Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice

Anders H. Lund[1]*, Geoffrey Turner[2]*, Alla Trubetskoy[2], Els Verhoeven[1], Ellen Wientjens[1], Danielle Hulsman[1], Robert Russell[3], Ronald A. DePinho[4], Jack Lenz[2] & Maarten van Lohuizen[1]

*These authors contributed equally to this work.

Published online: 19 August 2002, doi:10.1038/ng956

We have used large-scale insertional mutagenesis to identify functional landmarks relevant to cancer in the recently completed mouse genome sequence. We infected *Cdkn2a*$^{-/-}$ mice with Moloney murine leukemia virus (MoMuLV) to screen for loci that can participate in tumorigenesis in collaboration with loss of the *Cdkn2a*-encoded tumor suppressors p16INK4a and p19ARF. Insertional mutagenesis by the latent retrovirus was synergistic with loss of *Cdkn2a* expression, as indicated by a marked acceleration in the development of both myeloid and lymphoid tumors. We isolated 747 unique sequences flanking retroviral integration sites and mapped them against the mouse genome sequence databases from Celera and Ensembl. In addition to 17 insertions targeting gene loci known to be cancer-related, we identified a total of 37 new common insertion sites (CISs), of which 8 encode components of signaling pathways that are involved in cancer. The effectiveness of large-scale insertional mutagenesis in a sensitized genetic background is demonstrated by the preference for activation of MAP kinase signaling, collaborating with *Cdkn2a* loss in generating the lymphoid and myeloid tumors. Collectively, our results show that large-scale retroviral insertional mutagenesis in genetically predisposed mice is useful both as a system for identifying genes underlying cancer and as a genetic framework for the assignment of such genes to specific oncogenic pathways.

The Trp53 and Rb pathways are two of the principal pathways controlling cell proliferation that have been identified in human and mouse cells. The *Cdkn2a* locus is involved in both the p53 and Rb pathways by virtue of encoding p16INK4a, a regulator of CDK4/6-mediated Rb1 phosphorylation, and p19ARF, a modulator of Mdm2-mediated degradation of p53. Loss of *Cdkn2a* expression enhances the survival and proliferative potential of cells in the face of aberrant oncogenic signaling and cell-cycle entry[1]. The importance of *Cdkn2a* in tumor suppression is underscored by its frequent inactivation in human tumors by deletion, mutation or epigenetic silencing[2]. Mice deficient for both p16INK4a and p19ARF are viable but highly

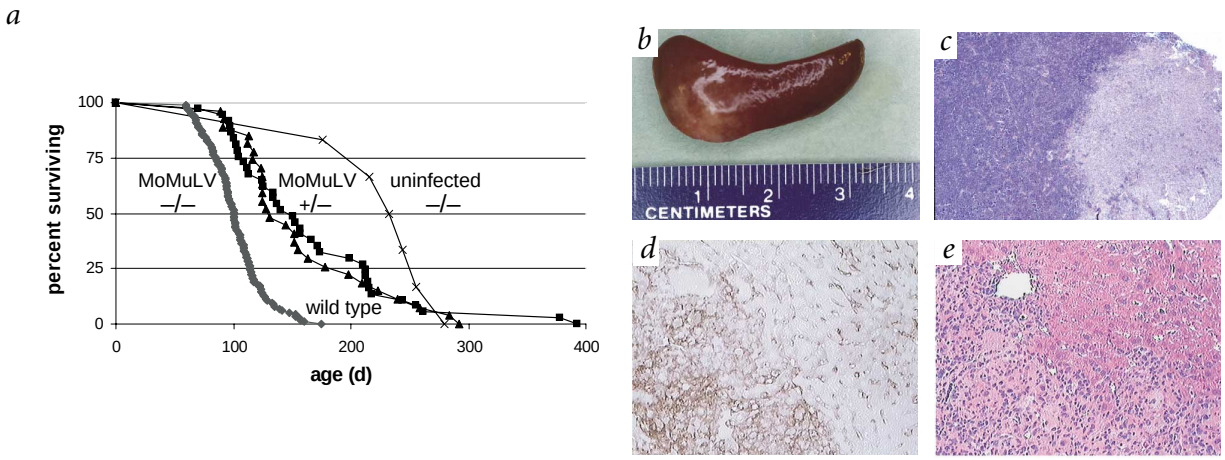

Fig. 1 MoMuLV-induced tumorigenesis in *Cdkn2a*$^{-/-}$ mice. ***a***, Disease-free survival of MoMuLV-infected *Cdkn2a*$^{-/-}$ mice after neonatal infection. Diamonds represent MoMuLV-infected *Cdkn2a*$^{-/-}$ mice (n = 115); triangles represent MoMuLV-infected *Cdkn2a*$^{-/+}$ mice (n = 37); squares represent MoMuLV-infected *Cdkn2a*$^{+/+}$ mice (n = 27); and 'x' marks represent uninfected *Cdkn2a*$^{-/-}$ mice (n = 6). We did not detect any evidence of a tumor in four of the infected *Cdkn2a*$^{-/-}$ mice. ***b***–***e***, Tumors in MoMuLV-infected *Cdkn2a*$^{-/-}$ mice. *b*, Spleen of a MoMuLV-infected *Cdkn2a*$^{-/-}$ mouse with a nodular tumor at the lower left. *c*, Hematoxylin and eosin staining of a spleen section of a MoMuLV-infected *Cdkn2a*$^{-/-}$ mouse with a histiocytic sarcoma, shown on the right at low magnification. *d*, The same section as *c* at higher magnification, stained with anti-F4/80 and horseradish peroxidase. Tumor cells uniformly stained with the antibody, whereas the staining in the normal portion of the liver is typical of that of hepatic mononuclear phagocytes[29]. *e*, Serial section from the same mouse as in *d*, stained with hematoxylin and eosin. This section shows the tumor at the lower left and a relatively normal hepatic architecture at the upper right.

[1]*Division of Molecular Genetics, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands. Departments of* [2]*Molecular Genetics and* [3]*Pathology, Albert Einstein College of Medicine, Bronx, New York, USA.* [4]*Department of Adult Oncology, Genetics and Medicine, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to J.L. (e-mail: lenz@aecom.yu.edu) or M.v.L. (e-mail: m.v.lohuizen@nki.nl).*

Table 1 • Integration frequencies into known common insertion loci

	Cdkn2a$^{-/-}$ tumors[a]		wt and +/– tumors	
Myc	10/51	20%	8/46	17%
Nmyc1	1/16	6%	2/23	9%
Pim1	5/51	10%	0/45	0%
Gfi1	14/51	27%	13/47	28%
Bmi1	0/51	0%	0/47	0%

[a]Frequencies of genomic rearrangements observed per number of tumors analyzed.

prone to tumors, succumbing to lymphomas and fibrosarcomas early in life[3]. Transcriptional regulation of *Cdkn2a* is complex, as evidenced by the growing number of proteins known to regulate its expression. These proteins include Bmi1, Dmp1, E2f family members, Ets, JunB, Myc and TBX2 (refs 4–10). Aberrant upstream signaling through Abl, β-catenin and Src and through the Ras–Raf–MAP kinase pathway can also effect activation of *Cdkn2a*[8,11–15].

To identify genes that act in collaboration with the combined loss of p16INK4a and p19ARF in tumorigenesis, we infected neonatal *Cdkn2a*$^{-/-}$ mice[3] with MoMuLV. This retrovirus induces tumors by insertional activation of proto-oncogenes or, more rarely, by insertional inactivation of tumor-suppressor genes. Although infection by MoMuLV elicits T-cell lymphomas, tumor susceptibility and phenotype is strain-dependent[16]. *Cdkn2a*$^{-/-}$ mice infected with MoMuLV developed tumors with a significantly shorter mean latent period than that of *Cdkn2a*$^{+/-}$ or *Cdkn2a*$^{+/+}$ littermate controls that were infected with MoMuLV ($P < 0.0001$, log-rank test) and that of uninfected *Cdkn2a*$^{-/-}$ siblings (Fig. 1*a*). This finding demonstrates a synergy between viral infection and loss of *Cdkn2a* expression in tumorigenesis.

We observed that 80% of *Cdkn2a*$^{-/-}$ mice infected with MoMuLV developed lymphomas. Approximately 60% of these mice had T-cell, 12% B-cell, and 8% mixed B- and T-cell lymphomas. In addition, approximately 55% had histiocytic sarcomas, which were distinctive nodular tumors (Fig. 1*b–c*), usually in the liver or spleen. These sarcomas stained positive for the macrophage marker F4/80

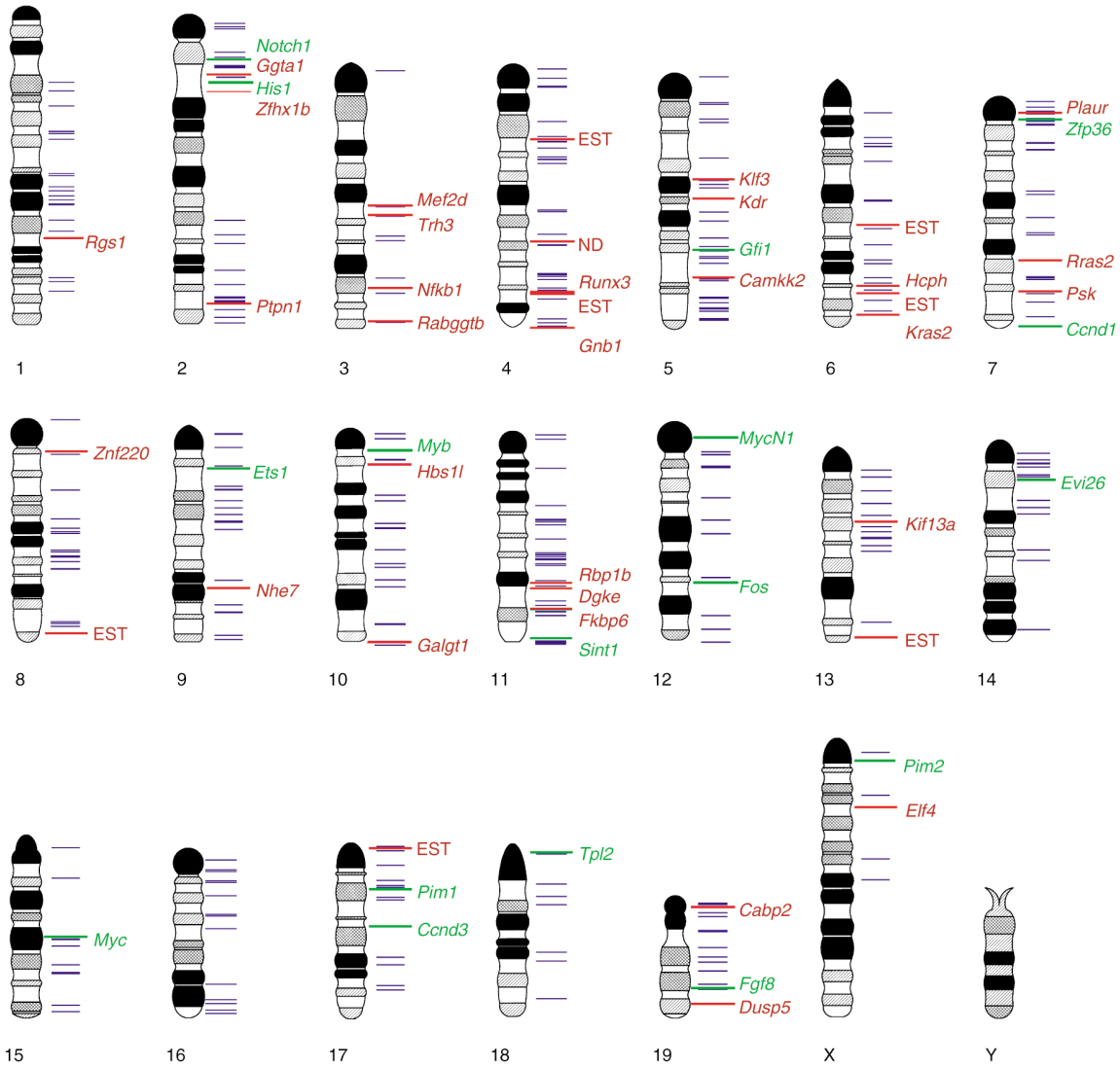

Fig. 2 Mapping of 565 unique retroviral insertion sites onto the mouse genome assembly from Celera Genomics. Blue lines indicate single MoMuLV insertions, green lines loci previously identified by retroviral insertional mutagenesis, and red lines newly discovered cancer-related CIS loci, with candidate gene names in red. Chromosome ideograms were modified with permission from C.V. Beechey and E.P. Evans (see Methods). Chromosome Y has not been assembled by Celera Genomics, and no tags were assigned to it.

(Fig. 1*d*) and were transplantable into immunodeficient SCID mice (data not shown). Moreover, 40% of MoMuLV-infected *Cdkn2a*$^{-/-}$ mice sustained both lymphomas and histiocytic sarcomas. Histiocytic sarcomas occurred in four uninfected *Cdkn2a*$^{-/-}$ mice, albeit at later times than in MoMuLV-infected mice. We observed histiocytic sarcomas in approximately 10% of *Cdkn2a*$^{+/-}$ mice infected with MoMuLV, but found none in infected wildtype controls, indicating that *Cdkn2a* deficiency or reduction predisposes to histiocytic sarcomas. All of the *Cdkn2a*$^{+/-}$ and *Cdkn2a*$^{+/+}$ mice infected with MoMuLV developed lymphomas, with no significant difference in the rate of tumorigenesis.

On the basis of the demonstrated synergy between MoMuLV insertional mutagenesis and loss of *Cdkn2a* function, we began a large-scale analysis of retroviral insertion site sequences to identify new oncogenes and tumor-suppressor genes. First, Southern-blot analysis of known CISs in MoMuLV-induced tumors revealed that genetic ablation of *Cdkn2a* did not significantly alter the insertion frequencies into these common loci (Table 1). Although these observations are based on a limited set of genes, it is likely that many specific tumorigenic events may benefit from loss of *Cdkn2a* expression, thus sensitizing the screen towards the identification of a broader spectrum of loci associated with cancer.

From a panel of 104 tumors—57 lymphoid (55%) and 47 myeloid (45%)—we isolated a total of 747 unique MoMuLV integration sites and directly sequenced them using a combination of inverse PCR and splinkerette-aided insertion site amplification. Homology searches in GenBank and in the nearly complete mouse genome databases from Celera Genomics and Ensembl allowed us to unambiguously map 565 viral insertions (Fig. 2). Several viral integrations were found to target the same locus in independent tumors, thus defining a CIS and indicating the presence of a cancer-associated gene[16]. The mapping data revealed that 163 (28.8%) of the sequences were clustered in 45 CISs; 37 of these have not been described previously (Table 2). In addition, we identified single viral integrations into many loci with known links to cell-cycle regulation or cancer pathogenesis, such as *Cdk2*, *Cdk8*, *Cdc25c*, *Mapk1*, *Fra1* and *E2f2*. Of the new CISs, 14 were specific to lymphomas, 6 were specific to the histiocytic sarcomas,

Table 2 • CISs identified in tumors of *Cdkn2a*$^{-/-}$ mice

Locus[a]	Celera ID[b]	Number of tags	Candidate gene[b]	GenBank accession number	Mouse chromosome	Ensembl mouse chromosome position[c]	Celera mouse chromosome position[c]	Human chromosome position	Cell type[d]
NKI-1	No ID	2	*Rgs1*	NM_015811	1	145163982–145186084	135762279–no hit	1q31	L
NKI-2	mCG21261	2	*Ggta1*	NM_010283	2	35932137–35932421	31638654–31639099	9q33-q34	L
NKI-3	mCG8151	3	*Zfhx1b*	NM_015753	2	45582724–45586359	41327870–41331182	2q22	L
NKI-4	mCG64382	8	*Ptpn1*	NM_011201	2	168789627–168845131	165422527–165471534	20q13.1–q13.2	L, HS
NKI-5	mCG8829	2	*Mef2d*	S68893	3	88925116–88929012	82005550–82008904	1q12–q23	L, HS
	mCG8841		*Iqgap1*	NM_016721				15q26	
NKI-6	mCG16756	3	RIKEN cDNA	NM_029789	3	95571261–95602603	87303050–87427968	1p21	L, HS
NKI-7	mCG22374	2	*Nfkb1*	NM_008689	3	106285108–no hit	127772992–127857753	4q24	L, HS
NKI-8	mCG7369	2	*Rabggtb*	NM_011231	3	154489997–154490618	146325593–146326261	1p31	L
NKI-9	mCG2332	2	putative	AK013175	4	43733132–43783247	41397143–41447334	ND	L, HS
NKI-10	mCG65233	3	RIKEN cDNA	XM_13153610	4	104505611–104665343	101393203–101543172	ND	L
NKI-11	mCG17461	3	*Runx3*	NM_019732	4	132639134–132789991	131734840–131779243	1p36	L, HS
NKI-12	mCG7764	2	RIKEN cDNA	NM_026383	4	133460472–133463348	132194843–132197692	ND	L, HS
NKI-13	mCG23363	2	*Gnb1*	NM_008142	4	150934253–151010713	152078608–152155328	1pter–p31.2	L
NKI-14	mCG10252	2	*Klf3*	NM_008453	5	63873382–63995956	59883104–60002945	4p16.1–p15.2	L, HS
NKI-15	mCG7209	2	*Kdr*	NM_010612	5	75154747–75155055	70951478–70952215	4q12	L, HS
NKI-16	mCG11116	3	*Camkk2,* mo	NM_031338	5	120807339–120868275	116383995–116512644	12q24.2	L, HS
NKI-17	mCG16286	2	KIAA0053, mo	NM_014882	6	88356371–88401020	86035121–86077929	ND	L, HS
NKI-18	mCG19794	2	*Hcph*	NM_013545	6	125626101–125626125	122564894–122564979	12p13	HS
NKI-19	mCG62286	2	EST	AV083176	6	130155495–130155708	127047268–127047481	ND	L
NKI-20	mCG13312	3	*Kras2*	NM_031515	6	145911824–145916127	140215453–140219542	12p12.1	L, HS
NKI-21	mCG7264	2	*Plaur*	NM_011113	7	16611803–16692064	9871320–9950546	19q13	L
NKI-22	mCG51745	2	*Rras2*	NM_025846	7	103758158–103810818	94601976– 94604902	11pter–p15.5	L
NKI-23	mCG22407	3	*PSK,* mo	NM_016151	7	116737933–116883522	112639555–112918319	16	L, HS
	mCG22413		*Maz*	AA915408				16p11.2	
NKI-24	mCG13090	2	*ZNF220* mouse ortholog	NM_006766	8	21537403–21537452	18309141–18309166	8p11	HS
NKI-25	mCG57225	3	EST	BB653745	8	126716162–126866935	120231285–120365173	ND	L, HS
NKI-26	mCG21612	2	*NHE7,* mouse ortholog	AF298591	9	95206300–95210881	90165786–90170216	X	L
NKI-27	mCG2827	6	*Hbs1l*	NM_019702	10	21080626–21195919	17946731–18022547	6q23–q24	L
NKI-28	mCG5175	2	*Galgt1*	NM_008080	10	127813456–127815198	109653266–109655579	12q13.3	L
NKI-29	mCG7666	6	*Rbp1B,* mo	AF199338	11	88413609–88417214	87602127–87605591	17q25	L, HS
NKI-30	mCG1480	2	*Dgke*	NM_019505	11	89769880–89818410	90822843–90872179	17q22	HS
NKI-31	mCG20534,	2	*Fkbp6*	NM_033571	11	101188993–101220645	102752458–102783766	7q11.23	L
NKI-32	mCG20453	2	*Kif13a*	AB037923	13	46365600–46409349	41811568–41852808	6p23	HS
NKI-33	mCG55520	2	EST	BF549796	13	110028533–110034445	101660853–101666913	ND	HS
NKI-34	mCG62490	2	EST	AV287890	17	3423327–3423365	891751–892712	ND	HS
NKI-35	mCG3890	4	*Cabp2*	NM_013878	19	6700612–6910493	2625769–2835944	11q13.1	L, HS
NKI-36	mCG20866	2	*DUSP5,* mo	NM_004419	19	53530620–5354765-	52164129–52180624	10q25	L
NKI-37	mCG5050	2	*Elf4*	NM_062654	X	34081336–34081401	33637118–33637542	Xq26	L, HS

[a]Newly discovered common insertion loci. Loci specific for the *Cdkn2a*$^{-/-}$ background, relative to previous screens[27,28], are indicated in bold. NKI, Netherlands Cancer Institute. [b]Candidate genes are based on their proximity to integrations. mo, mouse ortholog. [c]The position of both proximal and distal integration is indicated. [d]L, lymphoma; HS, histiocytic sarcoma; ND, not determined.

Table 3 • Viral insertions into known and new CISs

Locus	Number of tags	Mouse chromosome	Human chromosome
Known CISs			
Myc	22	15	8q24.12–q24.13
Gfi1	16	5	1p22
Tpl2	8	18	10p11.2
Pim1	7	17	6p21.2
Myb	4	10	6q22–q23
Ccnd1	3	7	11q13
Ccnd3	3	17	6p21
MycN1	3	12	2p24.1
Ets1	1	9	11q23.3
Evi26	1	14	10
Fgf8	1	19	10q24
Fos	1	12	14q24.3
His1	1	2	2q14–q21
Notch1	1	2	9q34.3
Sint1	1	11	17q25
Sox4	1	13	6p22.3
Zfp36	1	7	19q13.1
Newly discovered CISs[a]			
Dkmi5	1	3	1q21
Dkmi17	1	15	8q24
Dkmi22	1	19	11q13
Evi38	1	1	1q32
Evi42	1	2	9q34
Evi50	1	2	ND[b]
Evi62	1	4	1p36.13–p36.12
Evi73	1	5	7p22–p21
Evi78	1	6	2p21
Evi82	1	7	19
Evi92	1	8	4
Evi102	1	10	ND
Evi105	1	10	ND
Evi112	1	11	17q11.2
Evi130	1	17	6p21
Evi134	1	17	2p22.3–p21
Evi146	1	6	12pter–p13.31
Evi148	1	11	5q31.1

[a]Single insertions from *Cdkn2a*$^{-/-}$ tumors identified as CISs[27,28]. [b]Not determined.

and the remaining 17 consisted of sequences from both lymphomas and histiocytic sarcomas. As many mice had both histologically evident lymphoma and histiocytic sarcoma, it is likely that at least a fraction of the DNA samples were derived from tissue that was a mixture of both types of tumors.

Many of the newly discovered CISs encode proteins that can be assigned to functional pathways previously associated with mitogenic processes or transformation. The identification of two Ras homologs, *Kras2* and *Rras2*, is consistent with a large body of data establishing cooperative interactions between activated H-Ras and loss of Cdkn2a function[15]. Viral insertions near both *Gnb1* (at locus NKI-13) and *Gnb2* (single insertion) underline the importance of G-protein βγ-subunits in the regulation of signaling pathways involved in cell proliferation and survival[17]. Moreover, NKI-1, NKI-5 and NKI-17 also encode G-protein regulatory factors, and single viral insertions were mapped near *Rap1a*, *Rap1b* and an uncharacterized member of the Ras gene family. These findings further highlight the importance of deregulated Ras-like signaling in tumorigenesis. Three CISs, NKI-4, NKI-18 and NKI-36, target phosphatase genes. Both Ptpn1 and Hcph have been proposed to activate Src through removal of an inhibitory tyrosine phosphate within the Src-homology 2 (SH2) domain[18]. The CIS NKI-36 targets the mouse ortholog of the gene (*DUSP5*) encoding dual-specificity protein phosphatase, which is thought to negatively regulate the activity of the ERK1 MAP kinase[19]. Notably, the insertions in NKI-36 are found within the gene, indicating a potential tumor-suppressor function. The urokinase plasminogen activator receptor (NKI-21) is central to several pathways implicated in tumor progression. Aside from functions in angiogenesis and invasion, the urokinase plasminogen activator receptor activates the ERK1/2 MAP kinases, probably through interactions with β-integrins and receptor tyrosine kinases[20].

Transcription factors constituted a second large group of candidate oncogenes. The factor NFκB, identified in NKI-7, has important functions in cell-cycle progression, transformation and survival[21]. Myeloid elf-1 like factor (Elf4, NKI-37), a member of the Ets family, is a potent transcriptional activator in both lymphoid and myeloid cells. Notably, Elf4 recognizes consensus DNA-binding sequences similar to those recognized by other Ets proteins, which directly upregulate the *CDKN2A* promoter[7]. The protein ZNF220 is implicated in subtypes of acute myeloid leukemia (AML) carrying a t(8;16)(p11;p13) translocation[22]. The identification of the mouse ZNF220 ortholog as a CIS in histiocytic sarcomas (NKI-24) points to the use of this model to identify genes specifically involved in myeloid tumorigenesis. The targeting of a member of the AML gene family, *Runx3* (*Aml2*, NKI-11), is notable. Human *RUNX3* is located at 1p36, a region associated with structural aberrations in a number of neoplasias, including acute and chronic leukemia. In addition, RUNX3 was recently proposed to be an important tumor suppressor in human gastric cancers[23], raising the question of how Runx family members may exert both oncogenic and tumor-suppressor functions. In NKI-3, the integration hot spot targets *Zfhx1b*, encoding a transcriptional repressor recently shown to repress E-cadherin and promote tumor-cell migration[24].

The validity and sensitivity of the screen is evidenced by the identification of many genes previously implicated in MoMuLV-induced tumorigenesis (Table 3). Specifically, 22 integrations targeted *Myc*, a major collaborator in tumorigenesis with loss of *Cdkn2a*[25]. Reciprocally, no insertions among the 104 tumors analyzed targeted *Bmi1*, a known CIS previously shown to encode a repressor of *Cdkn2a*[4].

We confirmed the ability of the screen to identify collaborators of *Cdkn2a* loss using three independent approaches. First, Southern-blot analysis of a selected subset of the newly discovered CISs in tumors from *Cdkn2a*$^{-/-}$ and wildtype littermate controls revealed *Cdkn2a*$^{-/-}$-specific MoMuLV insertions in 2 of 45, 4 of 45 and 2 of 45 tumors analyzed for *Zfhx1b*, *Ptpn1* and *Rbp1B*, respectively. We found no viral insertions at these loci in a panel of 38 *Cdkn2a*$^{+/-}$ and *Cdkn2a*$^{+/+}$ tumors.

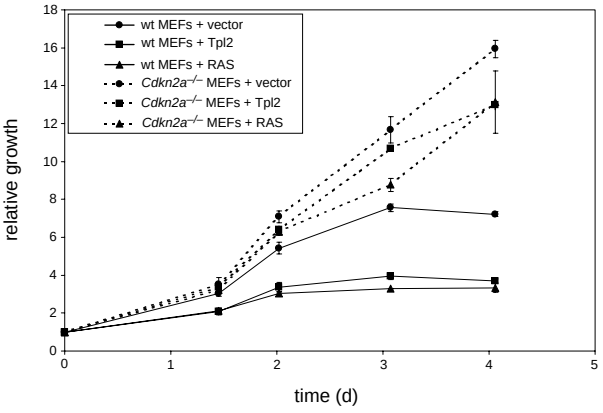

Fig. 3 Ectopic expression of Tpl2 collaborates with loss of *Cdkn2a*. The graph shows proliferation curves for wildtype (wt) and *Cdkn2a*$^{-/-}$ primary mouse embryo fibroblasts infected with pBabepuro-*Tpl2*, pBabepuro-*RASV12* or empty vector. Error bars indicate standard deviations for triplicate samples.

Second, we obtained experimental validation for a notable CIS within the screen: *Tpl2*, encoding a previously identified MAP kinase kinase kinase known to activate several MAP-kinase pathways[26]. Extending on the data generated from insertion site retrieval, Southern-blot analyses revealed MoMuLV insertions into *Tpl2* in 16 of 24 myeloid tumors analyzed, consistent with an important role of Tpl2 in myeloid tumorigenesis. As myeloid tumors arose exclusively in the *Cdkn2a*$^{-/-}$ background, no control myeloid tumors were available for analysis. Whereas five *Tpl2* insertions were found among tumors from mice sustaining mixed lymphoid and myeloid tumors or harboring sporadically appearing *Cdkn2a*$^{-/-}$ tumor types, no MoMuLV integrations were found at *Tpl2* among 47 *Cdkn2a*$^{-/-}$ and 30 *Cdkn2a*$^{+/+}$ or *Cdkn2a*$^{+/-}$ lymphomas. Ectopic expression of Tpl2 in mouse embryo fibroblasts (MEFs) resulted in diminished growth, resembling that seen in wildtype MEFs expressing oncogenic Ras. In contrast, the proliferation rate of *Cdkn2a*$^{-/-}$ MEFs infected with a *Tpl2*-expressing retroviral vector was similar to that of mock-infected cells (Fig. 3), and the cells eventually reached a higher cell density than did mock-infected *Cdkn2a*$^{-/-}$ MEFs (data not shown). These findings demonstrate a collaboration between *Tpl2* and loss of *Cdkn2a* and further support MAP-kinase activation as an important feature in *Cdkn2a* activation.

Third, to analyze the extent to which we have identified genes that act in collaboration with loss of the tumor suppressor Cdkn2a, we carried out extensive comparative data analyses against the retroviral tagging screens described in the accompanying papers by Suzuki *et al.*[27] and Mikkers *et al.*[28]. Of the 37 new CISs we identified, 18 loci were also identified as single or multiple hits in the screens described therein[27,28]. Thus, 19 new loci associated with cancer were identified on the *Cdkn2a*$^{-/-}$ background; these loci are targets for future investigations (Table 2). In addition, at several loci where we found overlaps among the three screens, we saw a significant skewing toward specific targeting of *Cdkn2a*$^{-/-}$. For example, we found nine insertions in *Ptpn1* on the *Cdkn2a*$^{-/-}$ background, whereas this locus is hit only once in each of the screens by Mikkers *et al.*[28] and Suzuki *et al.*[27] ($P < 0.001$, Fisher exact test). Similarly, targeting of the *Rbp1B* and *Zfhx1b* loci is significantly biased towards the *Cdkn2a*$^{-/-}$ background ($P < 0.02$ and $P < 0.05$, respectively). Although genetic predispositions, such as loss of *Cdkn2a*, favor activation of a specific subset of oncogenic pathways, the observed overlap between the three screens (including *Myc*, *Myb* and *Gfi1*) probably reflects common events in cancer that fall outside the functional groups identified in each of the systems used.

Single insertions from the *Cdkn2a*$^{-/-}$ screen overlapped 18 of the CISs identified by Mikkers *et al.*[28] or Suzuki *et al.*[27], raising the proportion of insertions within CISs to 36.4%. Combined CIS analysis of the three screens (where a CIS is defined as three insertions within a window of 100 kb) revealed 15 additional CIS loci[28]. This finding underscores the cumulative nature of large-scale insertional mutagenesis screens, the data from which have been assembled into a publicly available database.

Many of the candidate genes identified in this mouse model have been linked to cancer in humans or belong to gene families associated with human neoplastic diseases. Our identification of 37 new CISs, along with the large number identified in the accompanying papers[27,28], indicates that the ability to contribute to tumorigenesis is a potential property of many genes, not just a crucial few. Apart from its use in discovering new cancer genes, the power of the approach lies in the ability to focus on, and place these genes in, a limited set of signaling pathways associated with cancer. Owing to the specific genetic predisposition used in this screen, the majority of the loci can probably be functionally linked to the *Cdkn2a* locus, when the relevant candidate cancer genes have been verified, as is demonstrated by the preference for MAP-kinase activation via Tpl2. Evidence exists from both human tumors and rodent model systems showing that inactivation of *Cdkn2a* or downstream pathways is crucial in tumorigenesis. Thus, identification and functional characterization of genetic components acting on or in collaboration with these pathways is central to the understanding of many cancers.

Methods

MoMuLV tumorigenesis. *Cdkn2a*$^{-/-}$ and *Cdkn2a*$^{+/-}$ mice and wildtype littermates[3] were injected intraperitoneally with 10^5 infectious units of MoMuLV within 72 h of birth. We killed diseased mice, necropsied them and then carried out *Cdkn2a* genotyping by PCR. We froze most of each tumor for genetic studies. Fragments of tumors were fixed in either 10% formalin or Bouin's fixative, embedded in paraffin, sectioned and stained with hematoxylin and eosin. T-cell lymphomas appeared as lymphoblastic-like, diffuse tumors with pale-staining nuclei that were vesicular, pleiomorphic and frequently large. B-cell lymphomas appeared as lymphoblastic-like, diffuse tumors with moderately basophilic nuclei and prominent nucleoli. In some B-cell lymphomas, a fraction of cells showed immunoblastic morphology. We analyzed lymphomas by Southern blotting to assess T-cell receptor and immunoglobulin gene rearrangements (experimental details are available upon request). Cell-surface markers for hematopoietic lineages were determined by immunohistochemistry. We carried out histiocytic sarcoma transplantation into immunodeficient mice by subcutaneous injection, near the scapulae of BALB/c SCID mice, of approximately 4.5×10^5 disaggregated cells from macroscopically visible histiocytic liver nodules in virally infected *Cdkn2a*$^{-/-}$ mice. Tumors occurring in the livers of recipient mice histologically resembled those of donors. We confirmed that these tumors were of donor origin by PCR genotyping for the *Cdkn2a* allele with exon 2 deleted.

Retrieval and analysis of MoMuLV integration site sequences. We amplified 463 viral insertion sites using splinkerette-aided amplification essentially as described in Mikkers *et al.*[28]. We retrieved 284 insertion-site sequences using inverse PCR on *Sac*II-digested tumor DNA as described by Suzuki *et al.*[27]. All sequences have been deposited with GenBank (for exact insertion site positions, see Web Tables A–C online and the website listed below). Sequences of the MoMuLV–specific primers used for the nested inverse PCR are available upon request. We purified amplification products from agarose gels and subjected them to direct automated sequencing. To identify candidate genes, we filtered the insertion-site sequences for the presence of repetitive DNA and carried out homology searches using BLASTn in GenBank databases and in the mouse genome sequence databases from Celera Genomics (version 12) and Ensembl (Mouse Genome Sequencing Consortium, Mouse Assembly 3 February 2002 freeze).

Whereas CISs traditionally have been demonstrated using Southern blotting, the presence of the complete mouse genome allows for *in silico* CIS determination. Following a previously described statistical analysis[28], we defined retroviral common insertion sites as two or more integrations within 26 kb, or three insertions within 300 kb. For a set of 500 insertions, these windows give a tolerable statistically calculated background of approximately 2.5 CISs occurring at random. For insertions flanking a gene, the accepted distance between insertions was set to 100 kb. Although the functional distance between viral insertions and candidate oncogenes is known to differ between loci, the statistical threshold set here is in accordance with viral integration patterns surrounding previously characterized common insertion sites. In some loci, no candidate gene could be found in the immediate vicinity of an insertion. Moreover, a current limitation to the insertional mutagenesis approach is the limited extent and fidelity of the mouse genome annotation. This will improve over time. When analyzing overlaps between the *Cdkn2a*$^{-/-}$ screen and the screens carried out by Mikkers *et al.*[28] and Suzuki *et al.*[27], sequence tags falling within 100 kb on either side of a CIS were defined as overlapping.

Cell-culture assays. We cloned the *Tpl2* cDNA[26] into pBabe-puro. The retroviral vector pBabe-RASV12, procedures for growth, retroviral infection and proliferation analyses of primary mouse embryo fibroblasts have been described[25]. Briefly, MEFs were infected at passage 2, selected with 2 μg ml^{-1} of puromycin at passage 3 and seeded in triplicate with 25.000 cells per well in six-well plates at passage 4 for proliferation analysis. We measured relative proliferation using crystal violet staining. We repeated the experiment three times with independent batches of MEFs.

URLs. Data from the insertional mutagenesis screens are available at http://protagdb.nki.nl; Mo-MLV viral insertion site positions can be found at http://protagdb.nki.nl; the reference *Standard Ideogram/Anomaly Breakpoints of the Mouse* (Beechey, C.V. & Evans, E.P.) can be found at http://www.mgu.har.mrc.ac.uk/anomaly/anomaly-intro.html.

GenBank accession numbers. The retroviral flanking sequences have been deposited to GenBank, submission numbers AF524262–AF524826.

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

Acknowledgments

We thank L. Johnson, R. Stanley, D. Polsky, C. Cordon-Cardo, G. Hart, L. Spunk Jacobsen and R. Westrop for assistance. This work was supported by a grant from the National Institutes of Health (to J.L.) and a US National Institutes of Health training grant. A.H.L. was supported by a grant from the Danish Medical Research Council.

Competing interests statement

The authors declare that they have no competing financial interests.

Received 23 April; accepted 8 July 2002.

1. Sherr, C.J. The Cdkn2a network in tumour suppression. *Nature Rev. Mol. Cell Biol.* **2**, 731–737 (2001).
2. Ruas, M. & Peters, G. The p16INK4a/CDKN2A tumor suppressor and its relatives. *Biochim. Biophys. Acta* **1378**, F115–177 (1998).
3. Serrano, M. *et al*. Role of the INK4a locus in tumor suppression and cell mortality. *Cell* **85**, 27–37 (1996).
4. Jacobs, J.J., Kieboom, K., Marino, S., DePinho, R.A. & van Lohuizen, M. The oncogene and Polycomb-group gene *bmi-1* regulates cell proliferation and senescence through the ink4a locus. *Nature* **397**, 164–168 (1999).
5. Inoue, K., Roussel, M.F. & Sherr, C.J. Induction of ARF tumor suppressor gene expression and cell cycle arrest by transcription factor DMP1. *Proc. Natl Acad. Sci. USA* **96**, 3993–3998 (1999).
6. Bates, S. *et al.* p14ARF links the tumour suppressors RB and p53. *Nature* **395**, 124–125 (1998).
7. Ohtani, N. *et al.* Opposing effects of Ets and Id proteins on p16INK4a expression during cellular senescence. *Nature* **409**, 1067–1070 (2001).
8. Passegue, E. & Wagner, E.F. JunB suppresses cell proliferation by transcriptional activation of p16(INK4a) expression. *EMBO J.* **19**, 2969–2979 (2000).
9. Zindy, F. *et al.* Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. *Genes Dev.* **12**, 2424–2433 (1998).
10. Jacobs, J.J. *et al.* Senescence bypass screen identifies TBX2, which represses Cdkn2a (p19(ARF)) and is amplified in a subset of human breast cancers. *Nature Genet.* **26**, 291–299 (2000).
11. Damalas, A., Kahan, S., Shtutma -dependent growth arrest and cooperates with Ras in transformation. *EMBO J.* **20**, 4912–4922 (2001).
12. Zhu, J., Woods, D., McMahon, M. & Bishop, J.M. Senescence of human fibroblasts induced by oncogenic Raf. *Genes Dev.* **12**, 2997–3007 (1998).
13. Radfar, A., Unnikrishnan, I., Lee, H.W., DePinho, R.A. & Rosenberg, N. p19(Arf) induces p53-dependent apoptosis during abelson virus-mediated pre-B cell transformation. *Proc. Natl Acad. Sci. USA* **95**, 13194–13199 (1998).
14. Lin, A.W. *et al.* Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes Dev.* **12**, 3008–3019 (1998).
15. Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D. & Lowe, S.W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88**, 593–602 (1997).
16. Jonkers, J. & Berns, A. Retroviral insertional mutagenesis as a strategy to identify cancer genes. *Biochim. Biophys. Acta*. **1287**, 29–57 (1996).
17. Radhika, V. & Dhanasekaran, N. Transforming G proteins. *Oncogene* **20**, 1607–1614 (2001).
18. Bjorge, J.D., Jakymiw, A. & Fujita, D.J. Selected glimpses into the activation and function of Src kinase. *Oncogene* **19**, 5620–5035 (2000).
19. Ishibashi, T., Bottaro, D.P., Michieli, P., Kelley, C.A. & Aaronson, S.A. A novel dual specificity phosphatase induced by serum stimulation and heat shock. *J. Biol. Chem.* **269**, 29897–29902 (1994).
20. Aguirre Ghiso, J.A., Kovalski, K. & Ossowski, L. Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling. *J. Cell Biol.* **147**, 89–104 (1999).
21. Joyce, D. *et al.* NF-κB and cell-cycle regulation: the cyclin connection. *Cytokine Growth Factor Rev.* **12**, 73–90 (2001).
22. Borrow, J. *et al.* The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nature Genet.* **14**, 33–41 (1996).
23. Li, Q.-L. *et al.* Causal relationship between the loss of RUNX3 expression and gastric cancer. *Cell* **109**, 113–124 (2002).
24. Comijn, J. *et al.* The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol. Cell* **7**, 1267–1278 (2001).
25. Jacobs, J.J. *et al.* Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via Cdkn2a. *Genes Dev.* **13**, 2678–2690 (1999).
26. Patriotis, C., Makris, A., Chernoff, J. & Tsichlis, P.N. Tpl-2 acts in concert with Ras and Raf-1 to activate mitogen-activated protein kinase. *Proc. Natl Acad. Sci. USA* **91**, 9755–9759 (1994).
27. Suzuki, T. *et al.* New genes involved in cancer identified by retroviral tagging. *Nature Genet.* **32** 86–94 (2002).
28. Mikkers, H. *et al.* High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nature Genet.* **32** 73–79 (2002).
29. Roth, P., Dominguez, M.G. & Stanley, E.R. The effects of colony-stimulating factor-1 on the distribution of mononuclear phagocytes in the developing osteopetrotic mouse. *Blood* **91**, 3773–3783 (1998).