Sleeping beauty: a novel cancer gene discovery tool

Adam J. Dupuy, Nancy A. Jenkins and Neal G. Copeland*

Mouse Cancer Genetics Program, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA

Received February 1, 2006; Revised March 2, 2006; Accepted March 10, 2006

The National Cancer Institute and the National Human Genome Research Institute recently announced a 3-year 100-million-dollar pilot study to use large-scale resequencing of genes in human tumors to identify new cancer genes. The hope is that some of these genes can be used as drug targets for developing better therapeutics for treating cancer. Although this effort will identify new cancer genes, it could be made more efficient by preferentially resequencing genes identified as novel candidate cancer genes in animal models of cancer. Although retroviral insertional mutagenesis has proven to be an effective tool for identifying novel cancer genes in the mouse, these studies are limited by the fact that retroviral mutagenesis primarily induces hematopoietic and mammary cancer, but little else, while the majority of cancers affecting humans are solid tumors. Recently, two groups have shown that sleeping beauty (SB) transposon-based insertional mutagenesis can also identify novel candidate cancer genes in the mouse. Unlike retroviral infection, SB transposition can be controlled to mutagenize any target tissue and thus potentially induce many different kinds of cancer, including solid tumors. SB transposition in animal models of cancer could therefore greatly facilitate the identification of novel human cancer genes and the development of better cancer therapies.

INTRODUCTION

It has been over 20 years since the discovery of cellular oncogenes confirmed the genetic basis for cancer (1). Since this time the study of cancer genes has grown significantly in complexity. We now know that the majority of human cancer develops after a cell acquires loss-of-function mutations in tumor suppressor genes and gain-of-function mutations in oncogenes (2). Current estimates suggest that it may require as many as five or six such mutations acting together to transform the cell (3,4). In addition, so-called ‘epigenetic’ changes, alterations of gene expression because of chromatin modifications instead of mutation, are common in most human cancers. Overlapping but distinct sets of genes seem to be targeted by genetic and epigenetic alterations in cancer (5). Taken together these findings demonstrate the genetic complexity found in human cancer.

Recent advances in cancer therapeutics have shown that our knowledge of cancer genetics can be directly translated into improved cancer treatments. Herceptin (trastuzumab) is an antibody that binds to and inhibits the Her2 receptor encoded by the ERBB2 gene frequently amplified in breast cancer (6). Gleevec (imatinib) is a tyrosine kinase inhibitor that has shown to be effective in treating chronic myelogenous leukemia (CML) as well as gastrointestinal stromal tumors (GISTs), both of which are known to frequently harbor activating mutations in ABL and PDGFR tyrosine kinases (7,8). Finally, Iressa (gefitinib) is another tyrosine kinase inhibitor that blocks the activity of the EGF receptor often amplified in non-small-cell lung cancer (NSCLC) (9). These new targeted therapeutics have a range of activity from increased patient survival to clinical remission. Although these results are very encouraging, some patients do not respond to these drugs, and other responsive patients relapse after the tumor develops resistance to the drug (10–13). Nonetheless, these studies demonstrate the effectiveness of using targeted therapeutics in cancer treatment. These treatment strategies could be improved and new drugs developed through a better understanding of cancer genetics to identify better drug targets.

Despite advances in technology, the identification of cancer genes is still a slow and laborious process. Gene expression arrays allow for rapid assessment of transcript levels of thousands of genes simultaneously. Although gene expression profiling has made many correlative studies possible, it cannot be
used efficiently to identify mutations found within the genome of a tumor cell (14). Comparative genome hybridization (CGH) can be used to rapidly identify large-scale chromatin rearrangements (i.e. deletions, amplifications, etc.) in tumor cells but cannot be used to determine which gene or genes in these regions contribute to transformation (15). Recently, the Human Cancer Genome Project has been initiated to utilize large-scale genome sequencing to sequence the genomes of tumors in an effort to identify new cancer genes (http://www.genome.gov/Pages/About/NACHGR/May2005 NACHGRAgenda/ReportoftheWorkingGrouponBiomedical Technology.pdf). Although this effort will identify new cancer genes, it could be made more efficient by sequencing genes identified as candidate cancer genes in animal models of cancer. In addition, the Human Cancer Genome Project will likely identify many genetic alterations of undetermined significance. These will need to be complemented by reverse or forward genetic approaches.

IDENTIFICATION OF NOVEL CANCER GENES

The most common animal model used to study cancer is the mouse. Two general strategies have typically been used for identifying cancer genes in the mouse. The first involves generating engineered alleles (i.e. transgenic/knockout alleles) in the mouse germline to either inactivate a candidate tumor suppressor gene or introducing a transgene to overexpress a candidate oncogene to determine if specific mutations cause cancer. The second approach utilizes retroviral insertional mutagenesis to induce tumors in mice. Both strategies have been used successfully to identify new cancer genes, but each presents a unique set of obstacles.

Generating transgenic or knockout alleles is perhaps the most direct approach to determine the oncogenic potential of candidate genes. The validity of these models is a continuing source of concern, especially early models in which a candidate tumor suppressor was deleted in every cell of the mouse or a candidate oncogene was ubiquitously over-expressed (16). These situations are atypical in human cancer where tumor cells harboring mutations are surrounded by wild-type cells. This problem is beginning to be addressed through the use of the Cre/loxP system to generate spatially and temporally induced mutations (17). Even with these improvements, engineered alleles present several challenges. First, generating tumors in mice often requires two or more engineered mutations. This requires complex breeding schemes that make these experiments time-consuming and expensive. For this reason, engineered alleles are more often used to validate candidate cancer genes whose role in cancer is already implicated but rarely to identify novel cancer genes.

The second approach to inducing tumors in mice is more often used to identify novel cancer genes. Retroviruses such as the mouse mammary tumor virus (MMTV) and the murine leukemia virus (MuLV) do not encode genes that are acutely transforming. Instead these viruses introduce insertional mutations when they integrate into a host cell chromosome as a provirus. Integrated proviruses can introduce both gain-of-function and loss-of-function mutations thus activating oncogenes or inactivating tumor suppressor genes. These

mouse models of cancer are very useful in the identification of novel cancer genes because the proviruses not only induce mutations efficiently but also provide a sequence tag-marking the site of mutation within the tumor. In the post-genome sequence era these mutations can now be easily cloned using PCR-based approaches and mapped on the mouse genome. Hundreds of novel candidate cancer genes have been identified using insertional mutagenesis in mice (18,19). The major limitation with these models is that retroviral infection can produce mammary and hematopoietic cancer in mice but little else. Unfortunately, most cancer deaths are caused by solid cancers for which there are no insertional mutagen available to perform analogous screens in mice for novel cancer genes.

SLEEPING BEAUTY AWAKES WITH SOME EFFORT

Transposable elements are one type of insertional mutagen that has been exploited for decades by geneticists studying a variety of organisms (20,21). Sleeping Beauty (SB) became the first cut-and-paste transposon shown to have activity in mice (22). SB is a member of the Tc1/mariner family of transposable elements reconstructed from ancient elements found in the genomes of salmon fish. The system consists of two parts: the transposase enzyme and the DNA transposon vector. Initial work on the SB system in mice further demonstrated its activity, but transposition rates were below that needed for most genetic applications (23–26). Following these experiments several publications demonstrated improvements to the SB system in an effort to improve the efficiency of the system. The most significant factor affecting transposition rates appear to be the overall size of the transposon vector (27). Not surprisingly, transposons that are approximately the size of a wild-type Tc1/mariner element (~2 kb) hop at much higher rates than those of larger size. Additional experiments produced improved versions of both the transposon (28) and transposase (29,30) that improved transposition rates as well. However, despite these improvements, only modest increases in SB transposition rates in cell culture experiments were observed.

Two recent publications demonstrated for the first time that SB transposition can be used as a cancer gene discovery tool by making a few improvements to the SB system (31,32). First, two new mutagenic transposons were constructed specifically for the purpose of mutagenesis (T2/Onc, T2/ Onc2). These transposons carry a retroviral LTR and splice donor and can thus activate the expression of a proto-oncogene located nearby (Fig. 1A). In addition, these transposons carry splice acceptor sites on both DNA strands and a bidirectional polyA sequence and can thus disrupt the expression of a tumor suppressor gene when integrated in it (Fig. 1C and D). Likewise, the transposon can induce the expression of a truncated oncogene when integrated in it (Fig. 1B). These transposons were initially introduced into mice via a standard pronuclear injection in which a linear plasmid containing a single copy of the transposon forms a multicopy array at one site in the mouse genome (Fig. 2). These transposon donor arrays can be grouped according to
copy number with two transgenic lines having a moderate transposon copy number (31) and three having high copy number (32). As SB mobilizes transposons via a cut-and-paste mechanism, increasing the number of transposons in the cell should increase the number of transposition events. Two sources were used to drive transposase expression in these screens: a ubiquitously expressed transpose transgene (CAGGS-SB10) and a knock-in allele in which the transposase gene was knocked-in to the mouse Rosa26 locus (RosaSB) (23,32). Genes knocked-in to the Rosa26 locus are generally ubiquitously expressed in both embryonic and adult mouse tissues. Knock-in alleles are preferred to transgenes as they provide consistent, enforced gene expression. Transgenes, by comparison, are often silenced by direct methylation of the transgene thus producing unpredictable variegated patterns of gene expression.

The low copy transposon donor arrays together with the CAGGS-SB10 transposase transgene produced no spontaneous tumors in mice (31). In contrast, the high copy transposon donor array in combination with the RosaSB knock-in allele produced spontaneous tumors in mice (32). The prediction would be that the combinations of the low copy transposon array with RosaSB or the high copy transposon array with the CAGGS-SB10 transgene should produce an intermediate phenotype. In fact, the RosaSB transposase allele does produce spontaneous tumors in mice with low copy transposon arrays but at longer latency than with high copy donor arrays (David Largaespada, personal communication). In both cases, the RosaSB transposase allele produces lymphomas most frequently.

IDENTIFICATION OF TAGGED CANCER GenES USING TRANSPOSONS

As the transposons act as insertional mutagens they provide sequence tag-marking sites of mutation within the tumor. With recent advances in PCR technology and the availability of the mouse genome sequence, literally thousands of transposon integration sites can be cloned and mapped within a few weeks. This provides a broad picture of the mutations found within each tumor. This approach has been applied in two ways: to initiate and accelerate the development of tumors in wild-type mice or to accelerate the development of tumors in mice already predisposed to cancer.

In the first approach, transposons will provide the initiating mutations as well as the accelerating mutations. In this context, transposon mutagenesis can be used as a genetic tool to perform relatively unbiased forward genetic screens to identify not only novel cancer genes but also combinations of cancer genes that act together to transform a cell. For example, over 60% of the SB transposon-induced T-cell lymphomas described by Dupuy et al. (32) harbor activating transposon integrations in the Notch1 gene. Roughly, half of these T-cell lymphomas also harbor activating transposon integrations in the Rasgrp1 gene, a gene that functions upstream of Ras and activates Ras-signaling, suggesting that Ras-signaling cooperates with Notch1 signaling to produce these T-cell lymphomas. Identification of such cooperating cancer gene pathways will be extremely valuable for the development of improved targeted cancer therapeutics.

In the second approach, transposon mutagenesis can be used to identify cancer genes that cooperate with a specific initiating event such as loss of a tumor suppressor or activation of an oncogene. The complexity of integrations identified with this strategy will likely be decreased as each tumor cell carries a tumor initiating non-SB-induced mutation. However, it offers the ability to perform forward genetic screens for mutations that accelerate a defined mutation associated with
cancer. The first paper to use the SB system for this purpose mobilized transposons in mice deficient for the tumor suppressor p19<sup>Arf</sup> (31). Loss of p19 expression is associated with a number of human cancers, and mice lacking a functional p19 gene develop a wide variety of tumor types (33). In this context, the SB system could accelerate every tumor type that developed in p19 deficient mice. Analysis of 30 independent soft tissue sarcomas in these animals identified activating transposon integrations in the Braf locus as the most common mutation cooperating with p19 loss. In fact, the mutations caused by transposon integration in these tumors produce a truncated Braf protein with constitutive kinase activity. Point mutation of the Braf gene is thought to produce a similar constitutive kinase activity in human sarcoma, melanoma and other tumor types (34) validating the usefulness of SB for identifying cancer genes in solid tumors.

**CREATING CUSTOM CANCER MODELS IN MICE**

The available technologies for the identification of cancer genes involved in leukemia and lymphoma have been used successfully to identify hundreds of genes involved in these forms of cancer. Unfortunately, our current technical limitations make it much more challenging to identify cancer genes involved in most solid tumor types. The SB transposon system is one example of a new technology that could be used to identify candidate cancer genes in many forms of cancer. Regulated and/or inducible transposase alleles can be utilized to direct transposon-induced mutagenesis to specific sites in the mouse to model many forms of cancer. This could be done using a number of approaches. First, SB transposase could be expressed from a tissue-specific promoter instead of the ubiquitous Rosa26 locus (Fig. 3A). Investigators could take advantage of a number of cloned promoters or utilize BAC transgenes to target SB transposon expression to the site of interest.

A second approach would be to generate a Cre-inducible transposase allele. With this strategy, a ubiquitous SB transposase allele, such as the RosaSB allele, could be silenced by introducing a ‘lox-stop-lox’ cassette just upstream of the transposase cDNA (Fig. 3B). This approach has been used successfully to generate alleles that can be activated through the targeted expression of Cre recombinase (35,36). Investigators could then take advantage of a number of Cre alleles already available in mice to induce transposition at specific sites in the animal. In addition, Cre fusion proteins such as CreER could be used to initiate transposition at specific time points during embryonic development or in adult animals.

The transposon vector itself could also be modified, in addition to creating inducible and/or tissue-specific transposase alleles. Mutagenic transposons that lack an internal promoter could be used to exclusively disrupt gene expression in screens for novel tumor suppressor genes. Transposons containing insulator/silencing elements could be utilized to make broad alterations to chromatin structure often seen in human tumors. These are just a few possibilities, and many other modifications could be utilized as the transposon structure is quite flexible.

**TRANSLATING MOUSE CANCER INTO BETTER CANCER THERAPIES**

While cancer drugs designed against specific targets have shown much promise, clearly there is much work yet to be done. Recent advances in screening tools now make it possible to rapidly identify compounds with specific cellular targets (37). Unfortunately, our ability to identify the appropriate targets within tumor cells has limited the application of these tools. Mouse models of cancer generated using the SB transposon system could provide much needed insight into the complex genetic events that drive the initiation and progression of cancer. Initially, these models will be useful for the identification of candidate cancer genes in many forms of human cancer. Eventually, the SB system could be used to drive mutagenesis high enough in mice to link such clinically relevant processes as metastasis and drug resistance to specific gene mutations. It is also possible that transposon mutagenesis in mice will produce tumor models with sufficient genetic heterogeneity such that it could be used in preclinical testing of new cancer drugs. In this way, the SB system could contribute to many aspects of cancer research.

**ACKNOWLEDGEMENTS**

We thank David Largaespada for helpful comments on this manuscript. Adam J. Dupuy, Nancy A. Jenkins and Neal G. Copeland are supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

Conflict of Interest statement. None declared.
REFERENCES