The number of autosomal mammalian genes subject to random monoallelic expression has been limited to genes highly specific to the function of chemosensory neurons or lymphocytes, making this phenomenon difficult to address systematically. Here we demonstrate that asynchronous DNA replication can be used as a marker for the identification of novel genes with monoallelic expression and identify p120 catenin, a gene involved in cell adhesion, as belonging to this class. p120 is widely expressed; its presence in available cell lines allowed us to address quantitative aspects of monoallelic expression. We show that the epigenetic choice of active allele is clonally stable and that biallelic clones express p120 at twice the level of monoallelic clones. Unlike previous reports about genes of this type, we found that expression of p120 can be monoallelic in one cell type and strictly biallelic in another. We show that in human lymphoblasts, the silencing of one allele is incomplete. These unexpected properties are likely to be widespread, as we show that the Tlr4 gene shares them. Identification of monoallelic expression of a nearly ubiquitous gene indicates that this type of gene regulation is more common than previously thought. This has important implications for carcinogenesis and definition of cell identity.

Although the majority of mammalian genes are expressed from both parental alleles, there are some notable exceptions. In genomic imprinting, one allele is transcriptionally silenced, dependent on its parental origin (1). By contrast, X inactivation in females is a random process that leads to the silencing of most of the genes along one of the two copies of the X chromosome (2). This random choice is made independently by multiple cells around the time of implantation, and descendant cells maintain the choice, leading to mosaicism. In the last decade, a class of genes has emerged with properties similar to the genes subject to random X inactivation. Initially, the only known randomly monoallelically expressed genes on autosomes were the antigen receptors on lymphocytes (3), which were considered a special case because they undergo DNA rearrangement. Interest in this type of regulation was stimulated by the discovery that the members of a 1,000-gene family of olfactory receptor genes are also expressed in a random monoallelic fashion (4). Phorome receptors and a number of immune system molecules, including natural killer cell receptors and interleukins, have now been shown to be monoallelically expressed (5–8). Interleukins differ from other genes in this class in that their transcription can be monoallelic in some cells and biallelic in other cells (5, 7).

How can novel genes that are subject to random monoallelic expression be identified? Technical challenges of single-cell analysis of allelic choice in a mosaic tissue make identification of more genes with this type of regulation difficult. However, all classes of monoallelically expressed genes share the property of asynchronous DNA replication: during S-phase, one allele is replicated before the other. This DNA replication asynchrony has been assayed by fluorescence in situ hybridization (FISH) and S-phase fractionation methods (9–13). For genes subject to random monoallelic expression, asynchronous DNA replication is coordinated in a chromosome-wide fashion (13, 14). Importantly, the asynchrony is independent of expression of the gene in the assayed tissue (e.g. olfactory receptor gene replication is asynchronous in fibroblasts or ES cells, where these genes are not expressed). Thus, an attractive approach to identifying of novel loci is to detect asynchronous replication in a clonal cell line and then characterize expression of the candidate genes in the appropriate cell types.

Using this strategy, we found that the p120 catenin gene, which encodes a component of the cell adhesion machinery, is an asynchronously replicated gene. We show that it is monoallelically expressed in a subset of mouse and human clonal cell lines in a manner similar to interleukins. The p120 catenin gene is widely expressed (15), in contrast to the previously characterized monoallelically expressed genes. Expression of this gene in cell types that are easy to isolate and subclone allows for characterization of its transcriptional regulation.

MATERIALS AND METHODS

Animals and Cells—To obtain 129xCsF1 animals, 129/SvJ females were crossed with CAST/Ei males (Jackson Laboratory). Abelson cells were prepared from bone marrow of 6-week-old F1 animals or from embryonic day 14 liver by infecting primary culture with Abelson virus (16). Primary fibroblasts were generated simultaneously from ears of

* This work was supported in part by grants from the National Institutes of Health (to A. C.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

S The on-line version of this article (available at http://www.jbc.org) contains a supplemental list of BACs FISH probes, RFLP analyses, and primers.

† Supported by a Howard Hughes Medical Institute predoctoral fellowship.

¶ To whom correspondence should be addressed. Tel.: 617-258-7748; Fax: 617-258-6505; E-mail: chess@wi.mit.edu.
the same adult animals or from the embryonic tissue of the same embryo, respectively; after primary culture establishment, they were transformed with SV40. Cells were further cloned by limiting dilution (to the average of 0.3 cells/well) or by FACS. Normal human lymphoblastoid lines were obtained from Coriell cell repository, grown according to instructions, and cloned by FACS. Clonal ES cell lines used for primer extension replication timing analysis were provided by Dr. A. Wutz. EIA-transformed fibroblasts were provided by Dr. Elsa Flores. Replication Timing—FISH was performed essentially as described (10). Briefly, cells were given a 45-min pulse of BrdUrd before fixation. The DNA probe (BAC, cosmid, or PCR product as noted) was nick-translated (Amersham Biosciences kit) with Cy3-dCTP (Molecular probes) or FluorX-dCTP (Amersham Biosciences). After overnight hybridization and stringent washes, the nuclei were stained with anti-BrdUrd antibodies, stained with secondary antibody (Covance or Jackson ImmunoResearch Laboratories) coupled to fluorophore complementary to that of the probe, and counterstained with DAPI. For replication timing analysis, only BrdUrd-positive cells were counted. The complete list of probes is available in the supplemental material to this article.

Primer extension was performed as described previously (17) but with the following modifications. Ununsynchronized clonal populations of ES cells from 129xC571F1 mouse were grown in the presence of LIF. 10^8 cells were fixed in 75% EtOH in a standard procedure for cell cycle analysis (18), stained with DAPI in the presence of RNase A, and FACS-sorted into eight fractions: G_1, G_2, S_1–S_6 (left panel, typical FACS profile). DNA from each fraction was PCR-amplified with primers flanking a SNP distinguishing the 129 allele of an assayed gene from the Cs allele. Products of allele-specific primer extension were detected using a mass spectrometer (center). Un, mass peak of unextended primer. To calibrate for bias and confirm linearity of allele detection, each experiment included known mixes of parental DNA (right), mean ± S.E., n = 4. B, primer extension assay readout. For each of the eight cell cycle fractions, the relative content of 129 (maternal) allele is presented as mean ± S.E. (n = 4). No significant difference from 1:1 ratio of each allele marks synchronously replicated gene Adam1a (left panel), whereas for asynchronously replicated genes, such as olfactory receptor mORT2 (center), an overabundance of the early replicated allele causes relative allele content in some S-phase fractions to differ significantly from 50%. p120 (right panel) was paternally early. C, FISH-based assay for asynchronous replication. Mouse fibroblast nucleus with an SD pattern is shown. Red depicts 10-kb PCR probe for p120 labeled with Cy3, and blue depicts DAPI. D, summary of the FISH assay in primary mouse fibroblasts with probes for p120 and for the asynchronously replicated imprinted gene H19, the randomly asynchronously replicated gene IL-4, and a synchronously replicated locus Ccnf. SS, single dot-single dot; DD, double dot-double dot. In each case, n = 50–100. Similar results were observed with both BAC and PCR probes for p120 as well as in ES cells, EIA-transformed fibroblasts, and Abelson cells.

Asynchronous Replication of p120 Catenin—We identified the p120 catenin gene as an asynchronously replicating gene (Fig. 1, A and B). The first approach involved S-phase fractionation, PCR, and primer extension to analyze allele-specific DNA replication in cells derived from an F1 mouse as described previously (13, 17). We confirmed the asynchronous replication of p120 catenin using a standard FISH assay (4, 10, 13) in primary mouse fibroblasts (Fig. 1, C and D) as well as SV40 and EIA-transformed fibroblasts, ES cells, and Abelson cells (not shown). In this FISH analysis, replicated loci are visualized as a double dot hybridization signal, whereas unreplicated loci reveal a single dot. Asynchronously replicating genes reveal a single dot-double dot (SD) pattern in 30–40% of S-phase cells as opposed to around 10–20% SD for synchronously replicating genes.

Monoallelic Expression of Mouse p120 Catenin—We analyzed p120 catenin expression in Abelson murine leukemia virus (Abl-MLV)-transformed pre-B cell clonal cell lines derived from an F1 cross between a female 129 mouse and a
male *Mus musculus castaneus* (Cs) mouse. Clonal cell lines were made by using a FACS to place a single-cell well or by limiting dilution. Both approaches gave similar results. After determining that these cell lines express all four of the splice variants described previously, we proceeded to assess the relative mRNA levels of the two alleles of the p120 gene. We took advantage of an SNP present in the 3′ untranslated region of the message that can be distinguished using either primer extension or a polymorphic HaeIII restriction endonuclease site (RFLP analysis). Examples of restriction endonuclease analyses revealing monallelic and biallelic cell lines are shown in Fig. 2A. In total, we analyzed 78 clonal Abl-MLV-transformed pre-B cell lines (derived from eight mice) for p120 expression using RFLP. Analyses of allele-specific expression were also performed using Sequenom mass spectrometric primer extension genotyping of RT-PCR products (Fig. 2D). Although this genotyping platform is normally used to call heterozygotes and homozygotes, it has been demonstrated to detect the presence of a rare allele at 2% in a complex mixture (23, 24). Therefore, for the Abl-MLV-transformed lines with only one allele detected, there is at least 50-fold skewing. Sixty percent of the lines monallelically express the p120 catenin gene (32% 129 and 28% Cs), with the remaining 40% of the lines showing biallelic expression. This distribution of monallelic versus biallelic expression is similar to what has been observed for the interleukin-2 and -4 genes (5, 7).

The stability of the allele-specific expression pattern was assessed both for monallelic and biallelic clonal cell lines. Monoallelic lines were stable for up to 50 weeks maintained in continuous culture (Fig. 2B). Biallelic clonal cell lines were also stable, and in one case, we performed subcloning of a bialleic line and analyzed 15 subclones, all of which maintained biallelic expression qualitatively similar to the parental clonal cell line. Abl-MLV-transformed cell lines express all four reported splice variants of p120 catenin (15) regardless of whether the paternal allele, maternal allele, or both alleles are expressed (Fig. 2C).

To assess the relative levels of transcription in monallelic versus biallelic lines, we employed real-time quantitative PCR analysis of RT products, in each case normalizing the p120 level to the measured level of the glyceraldehyde-3-phosphate dehydrogenase transcript. On average, the biallelic clones express p120 catenin at twice the level of monoa llelic clones (Fig. 2E). Although there is a relatively high variation in the level of expression among both monoa llelic and biallelic cell lines, a Student’s *t* test analysis (two-tailed with unequal variance) indicates that the difference between the two distributions is significant at the *p* < 0.05 level. The distributions are consistent with a 2-fold difference in expression between monoa llelic and biallelic expression, suggesting that if a given allele is transcribed, its level of expression is independent of whether or not the other allele is active.

**Asynchronous Replication and Monoallelic Expression of the Human p120 Catenin Gene**—Replication timing of the human p120 catenin gene was examined using the FISH-based replication timing assay on primary human fibroblasts. The human p120 gene reveals 41% of S-phase cells with an SD pattern, which is consistent with other asynchronously replicating genes (Fig. 3A and B). Thus, as was observed for the mouse gene, the human p120 catenin gene is asynchronously replicating.

The relative expression levels of the two alleles of the human p120 gene were determined by analyzing clonal Epstein-Barr virus-transformed B cell lines derived from an individual heterozygous for a BstNI RFLP in the p120 catenin cDNA. Subclones were generated using single-cell FACS sorting and subjected to RT-PCR after sufficient cell expansion. Digestion of the RT-PCR product with the BstNI endonuclease reveals that one allele is digested and the other allele is not digested.

Analyses of 23 single-cell-derived subclones show a variety of patterns (Fig. 3C) with some cell lines preferentially expressing one allele and others preferentially expressing the other allele. Other cell lines express roughly equal levels of both alleles. A control analysis of genomic DNA derived from the parental cell line indicates that the primers amplify both alleles with a slight bias toward the undigested allele. The extent of allele-specific bias of the expression of the human p120 gene is not as great as the extent of skewing observed in the mouse. We estimate the extent of the human gene bias to be 10–20-fold in some lines (Fig. 3C). By contrast, in the murine Abl-MLV-
Monoallelic Expression of p120 Catenin

We have demonstrated a new strategy for identifying a novel gene with random monoallelic expression. We detected asynchronous replication of p120 catenin, confirmed it by an inde-

transferred clones monoallelically expressing p120, the skewing was at least 50-fold if not absolute (Fig. 2D).

**Monoallelic Expression of the Mouse Tlr4 Gene**—To examine a second monoallelically expressed gene in the same murine Ab1-MLV-transformed clones analyzed above, we examined the Tlr4 receptor gene, which is monoallelically expressed in B cells (25). The Tlr4 receptor is involved in innate immunity. We demonstrated that Tlr4 is asynchronously replicating in mouse cells (Fig. 4, A and B). Analyses of clonal cell lines indicated that one-fourth have monoallelic expression for either the maternal or paternal allele (Fig. 4C). However, these cell lines with monoallelic expression are not absolute in their monoallelic expression. Rather, we observe skewing of expression reminiscent of the skewing observed in analysis of the human p120 catenin gene. Based on control PCR experiments with known mixes of the two alleles as input, we estimate that the skewing observed in some Ab1-MLV-transformed B cell lines is ~20-fold (data not shown). Recall that the skewing observed for the human p120 catenin gene is ~10-fold. The mouse Tlr4 is therefore a second example of an autosomal randomly monoallelically expressed gene that is not absolute in its allele-specific transcription patterns. We also examined clonal fibroblast cell lines for allele-specific transcription of Tlr4 and found only biallelic lines (12 lines). Note that similar analyses of 14 clonal fibroblast lines revealed only lines biallelic for expression of the p120 catenin gene.

Tlr4 and p120 catenin are on different chromosomes; therefore, based on our prior observation that asynchronous replication is independently regulated for different chromosomes, we asked whether the allele specificity of monoallelic transcription of Tlr4 and p120 catenin are also independently regulated. We addressed this question by analyzing clonal Ab1-MLV-transformed B cell lines for allele-specific expression of both p120 catenin and Tlr4. In total, 43 lines were examined both for the Tlr4 and p120 catenin genes. Six of the lines were monoallelic for both genes, nine were biallelic for both genes, and 28 were monoallelic for one gene and biallelic for the other. These analyses reveal that the two genes are independent in their monoallelic expression (Fig. 4C). In the four examples shown, clonal Ab1-MLV-transformed lines express p120 catenin from the maternal allele. The maternal and paternal alleles of p120 can be distinguished using an RFLP; the maternal (129 allele) is digested by the AflIII restriction endonuclease. One of these lines predominantly expresses the parental Tlr4 allele, one line predominantly expresses the maternal allele, and two lines express equal amounts of the two alleles. We also analyzed clonal cell lines prepared by limiting dilution, which also reveal monoallelic expression for both genes (Fig. 4C). Thus, in terms of both whether there is monoallelic expression and which parental allele is monoallelically transcribed, these two genes are independently regulated.

**DISCUSSION**

We have demonstrated a new strategy for identifying a novel gene with random monoallelic expression. We detected asynchronous replication of p120 catenin, confirmed it by an inde-
dependent method, and then characterized the allele-specific expression of the gene in fibroblasts and cells of the B cell lineage. The advantage of this approach is that it does not depend on whether the candidate gene is expressed in the cells used for initial analysis. This should allow for the scaling up of this approach to identify candidate genes in a systematic fashion, including genes expressed in an extremely restricted manner.

Our observation of allele-specific transcription of the p120 catenin gene has added several unexpected facets to our understanding of the epigenetic phenomenon of random monoallelic expression. Unlike the other genes in this class, such as olfactory receptors and T and B cell receptors, p120 is very widely expressed in human and mouse tissues. This expands the class of such genes in a crucial way, suggesting that other widely expressed genes may share this mode of transcriptional regulation. p120 catenin is similar to interleukins in that although a majority of Abelson cells express it from a single allele, a significant fraction of cells is expressing both alleles equally.

We found that p120 catenin is asynchronously replicated and monoallelically expressed in both mouse and human cell lines. This suggests that natural selection favored the features responsible for this unusual type of gene regulation over the ~60 million years separating primates and rodents from their last common ancestor. Monoallelic expression of p120 may be directly advantageous for mammals. It is also possible that it is a consequence of the genome architecture in this region, because in both mouse and human genome, p120 is less than 1 megabase away from clusters of olfactory receptors.

We observed, for the first time, that a gene can be randomly monoallelically expressed in one cell type (p120 in Abelson cells) and completely biallelically expressed in another (fibroblasts, Fig. 2). In practical terms, this shows that observing biallelic or monoallelic expression of a given gene in one cell type is not necessarily predictive for other cell types. More generally, this observation raises an important mechanistic question: what is the difference between the cells that are monoallelic or biallelic with respect to a given gene? Either their DNA carries different epigenetic modifications or the modifications are identical, but cells of one type are competent for transcriptional readout of such modifications and cells of another type are not. Note that some epigenetic mark is always present, as asynchronous DNA replication of p120 is ubiquitous. Further study of p120 and similar genes will help to distinguish these possibilities.

The question of causal relationship between asynchronous DNA replication and random monoallelic expression is very intriguing. To our knowledge, all known genes subject to random monoallelic expression are replicated asynchronously, an observation further confirmed by asynchronous replication of Tlr4 locus (Fig. 4). Asynchronous replication thus appears to be a necessary condition of random monoallelic expression. It is, however, not a sufficient condition. For example, transcription of p120 is completely biallelic in fibroblasts, even as its replication is asynchronous in these cells. Moreover, further analysis should determine whether early or late replication status of a p120 allele determines the activity of the allele. We have shown previously that for Igκ-κ rearrangement occurs on the early replicating allele, rendering it active (12); this may also hold true for non-rearranged loci.

Determining the relative allele activity is an inherent challenge in the analysis of random monoallelic expression. Expression in tissues is mosaic, and single-cell approaches provide limited ability for quantitative study. The fact that p120 is transcribed from either paternal or maternal allele in stable clonal cell lines allowed us to analyze its expression in more detail. The quantitative analysis of p120 expression in human lymphoblastoid lines revealed the unexpected feature of incomplete silencing, where in some cell lines, one allele is about 10-fold less active than the other (Fig. 3). This resembles the incomplete and variegated silencing of X-linked gene Rep1 in female cells (26).

Another intriguing observation is that cell lines expressing p120 from both alleles have a level of p120 transcript about 2-fold higher than in the lines expressing p120 from a single allele (Fig. 2E). Thus, similar to X inactivation, it appears that the transcriptional regulation is centered not on the total level of transcript in the cell but on each given promoter. When the second allele is “on,” both alleles are equally active, resulting in a double dose of the transcript. Therefore, monoallelic expression in a fraction of cells of a given type creates variability in the dose of expression of the affected gene. It is especially noteworthy that even with two absolutely identical alleles, the difference in the transcript level of the affected gene between cells expressing the gene from one allele and cells expressing it from both alleles can cause these cells to be functionally different. This is potentially relevant to tumorigenesis, as discussed below.

Both incomplete silencing and cell type dependence of monoallelic expression are not unique features of p120, as we show that Tlr4 shares these properties in mouse cells (Fig. 4), suggesting that these properties may be relatively widespread. Another point arising from our analysis of monoallelic expression of Tlr4 and p120 in the same cells is that these genes are independently regulated. We have shown previously that the epigenetic mark responsible for asynchronous replication is independent for different autosomes, even as it is coordinated on a given autosomal pair (13). Considering that p120 and Tlr4 are on different chromosomes, it is thus not surprising that their transcriptional choices are independent of one another. However, an important point arises with monoallelic genes that could be coexpressed in the same cell (similar to p120 and Tlr4). Even the cells with exactly the same expression profile would differ in their complement of alleles expressed from multiple genes of this type. The combinatorial possibilities are large and add an extra dimension to the definition of cell identity.

An especially intriguing hypothesis concerns variability in allele-specific expression in genes involved in carcinogenesis. Increased epigenetic heterogeneity has been suggested as a contributing factor in metastatic potential of tumor cells (27). The exclusive expression of a defective allele of a tumor suppressor gene would increase chances of malignant transformation in this cell (28). A more subtle possibility is presented by p120 catenin. As part of the E-cadherin complex (29), p120 catenin interacts with tyrosine kinases (30, 31), and their activity impacts the strength of cell-cell adhesion. Overrepresentation of individual splice variants of human p120 catenin has been shown to dramatically affect cell shape (32), and the observation of changes in expression of p120 in tumors has led to the suggestion of a role for p120 in tumorigenesis and/or metastasis (29, 33–35). Monoallelic expression could cause even a subtle mutation (e.g. a mutation affecting preferential outcomes of alternative splicing) to change the function of the cell expressing only the mutated allele. Moreover, as noted above, even with two identical alleles, random monoallelic expression can cause variation in levels of expression, including that of p120. These ideas raise an interesting possible connection between monoallelic expression that is present in normal cells and early events in tumorigenesis.

Acknowledgments—We thank S. Gabriel, M. Defelice, and B. Blumenstiel for assistance with Sequenom genotyping system;
G. Paradis for help with FACS sorting; C. Cowles for advice with primer-extension genotyping; E. Winchester, H. Skaletsky, G. Bell, and B. Yuan for help with software; H. Higgins for help in preparing the manuscript; and R. Jaenisch, T. Orr-Weaver, A. Bortvin, and members of the Chess laboratory for useful discussions.

REFERENCES

1. Reik, W., and Walter, J. (2001) Nat. Rev. Genet. 2, 21–32
2. Lyon, M. F. (1986) Nature 320, 31–32
3. Pernis, B., Chiappino, G., Kelus, A. S., and Gell, P. G. (1965) J. Exp. Med. 122, 853–876
4. Chess, A., Simon, I., Cedar, H., and Axel, R. (1994) Cell 78, 823–834
5. Bix, M., and Locksley, R. M. (1998) Science 281, 1352–1354
6. Hollander, G. A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D. E., Bhan, A. K., and Burakoff, S. J. (1998) Science 279, 2118–2121
7. Rhoades, K. L., Singh, N., Qi, P., Gribnau, J., and Chess, A. (2000) Curr. Biol. 10, 455–463
8. Held, W., Roland, J., and Raulet, D. H. (1995) Nature 376, 355–358
9. Taylor, J. H. (1960) J. Biophys. Biochem. Cytol. 7, 455–463
10. Mostoslavsky, R., Singh, N., Tenzen, T., Goldmit, M., Gabay, C., Elizur, S., Qi, P., Reubinoff, B. E., Chess, A., Cedar, H., and Bergman, Y. (2001) Nature 414, 221–225
11. Kitsberg, D., Selig, S., Brandeis, M., Simon, I., Keshet, I., Driscoll, D. J., Nicholls, R. D., and Cedar, H. (1993) Nature 364, 459–463
12. Anastasiadis, P. Z., and Reynolds, A. B. (2001) Curr. Opin. Cell Biol. 13, 694–610
13. Thoreson, M. A., and Reynolds, A. B. (2002) Differentiation 70, 583–589
14. Mayerle, J., Friess, H., Buchler, M. W., Schnekenburger, J., Weiss, F. U., Zimmer, K. P., Domschke, W., and Lerch, M. (2003) Gastroenterology 124, 949–960
15. Latt, S. A. (1973) Proc. Natl. Acad. Sci. U. S. A. 70, 3395–3399
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
17. Ohlsson, R., Tycko, B., and Sapienza, C. (1998) Trends Genet. 14, 435–438