Monoallelic Expression of Pax5: A Paradigm for the Haploinsufficiency of Mammalian Pax Genes?

Stephen L. Nutt and Meinrad Busslinger1,*
Research Institute of Molecular Pathology,
Dr. Bohr-Gasse 7, A-1030 Vienna, Austria

* Corresponding author

It is generally assumed that most mammalian genes are transcribed from both alleles. Hence, the diploid state of the genome offers the advantage that a loss-of-function mutation in one allele can be compensated for by the remaining wild-type allele of the same gene. Indeed, the vast majority of human disease syndromes and engineered mutations in the mouse genome are recessive, indicating that recessiveness is the ‘default’ state. However, a minority of genes are semidominant, as heterozygous loss-of-function mutation in these genes leads to phenotypic abnormalities. This condition, known as haploinsufficiency, has been described for five of the nine mammalian Pax genes, which are associated with mouse developmental mutants and human disease syndromes. Recently we have reported that the Pax5 gene is subject to allele-specific regulation during B cell development. Pax5 is predominantly transcribed from only one of its two alleles in early B-lymphoid progenitors and mature B cells, while it transiently switches to a biallelic mode of transcription in pre-B and immature B cells. As a consequence, B-lymphoid tissues are mosaic with regard to the transcribed allele, and heterozygous mutation of Pax5 therefore results in deletion of B lymphocytes expressing only the mutant allele. The allele-specific regulation of Pax5 raises the intriguing possibility that monoallelic expression may also be the mechanism causing the haploinsufficiency of other Pax genes. In this review, we discuss different models accounting for the haploinsufficiency of mammalian Pax genes, provide further evidence in support of the allele-specific regulation of Pax5 and discuss the implication of these findings in the context of the recent literature describing the stochastic and monoallelic activation of other hematopoietic genes.

Key words: B cell development / Haploinsufficiency / Monoallelic expression / Pax5 / Single-cell RT-PCR.

Models to Account for Haploinsufficiency

The genetic term haploinsufficiency describes a situation where a null mutation in one allele of a gene results in a mutant phenotype despite the presence of the second, wild-type allele. A haploinsufficient gene therefore fulfills its normal function only in the presence of both wild-type alleles, and hence its activity is highly sensitive to gene dosage.

Haploinsufficiency is a relatively rare genetic phenomenon in mammals, as the majority of human disease syndromes and engineered mouse mutations are recessive (Fisher and Scambler, 1994). Nevertheless, a growing number of mammalian transcription factor genes have recently been reported to be haploinsufficient (Fisher and Scambler, 1994; Engelkamp and van Heyningen, 1996). While several explanations for haploinsufficiency have been suggested (Wilkie, 1994; Read, 1995), we will discuss below only the three most likely models.

The Protein-Protein Interaction Model

Transcription factors commonly depend on the interaction with partner proteins to fulfill their functions in gene activation and repression. Both positive and negative control by interacting proteins is frequently observed for transcription factors which bind DNA as heterodimeric complexes. This type of regulation is best illustrated by the family of transcription factors which interact with each other through the basic helix-loop-helix (bHLH) domain. As the function of these regulators is determined by the relative ratio of activating and inhibitory HLH proteins (Bain and Murre, 1998), it is conceivable that a two-fold change in the concentration of one partner protein could have a dramatic effect on the regulation of target genes within the cell. Indeed, several bHLH transcription factors including Sisterless-b (Sis-b) play an essential role in Drosophila sex determination where they cooperate to amplify a small quantitative (two-fold) difference in sis gene dose into a qualitative all-or-none effect on the transcription of the target gene Sex-lethal (Cline, 1993; Parkhurst and Meneely, 1994). A second example was revealed by targeted mutation of the murine E2A gene which demonstrated an essential role for this bHLH transcription factor in B cell development (Bain et al., 1994; Zhuang et al., 1994). The E2A mutation appears to be haploinsufficient, as 50% less B lymphocytes were detected in the fetal liver of heterozygotes compared to wild-type embryos (Zhuang et al., 1994). Moreover, increased expression of...
the inhibitory HLH protein Id1 in the B-lymphoid lineage resulted in a phenotype resembling the E2A null mutation (Sun, 1994), indicating that B-lymphopoiesis is extremely sensitive to changes in the relative ratio of positively and negatively acting HLH proteins.

**Binding Site Occupancy Model**

This hypothesis is based on the assumption that the genome contains a large number of DNA-binding sites for a particular transcription factor which may itself be present only in limiting amounts within the cell. Under these conditions, only the binding sites of relatively high affinity will be occupied and hence potentially regulated by the transcription factor (Read, 1995). A small increase or decrease in protein concentration could consequently lead to a change in binding site occupancy and thus to an altered profile of target gene expression. Evidence in support of this model was obtained by analysis of the paired domain-containing transcription factor Pax6 (Schedl et al., 1996). Pax6 is known to be haploinsufficient in mammals, as heterozygous mutations in this gene are responsible for the Small eye phenotype in mouse and for congenital eye abnormalities known as aniridia and Peters’ anomaly in humans (Hanson and van Heyningen, 1995). Surprisingly, a modest increase in Pax6 expression due to the presence of additional copies of the wild-type Pax6 locus generates a phenotype which is similar to that of the heterozygous Small eye mouse (Schedl et al., 1996). Hence, the transcription factor Pax6 is exquisitely dosage-sensitive, as it can control normal eye development only within a narrow concentration range.

**The Monoallelic Expression Model**

A third possible mechanism to account for haploinsufficiency is the monoallelic expression theory, which we will describe in detail in this review. According to this hypothesis, a given gene is transcribed from only one of its two alleles which are selected for expression on a purely stochastic basis. Consequently, the expressing tissue is mosaic with regard to transcription of this gene and can be subdivided into two cell populations, each transcribing either allele 1 or 2, respectively (Figure 1). Both cell populations can participate in normal development and differentiation, as long as the two alleles contain wild-type sequences. However, if one allele is mutated, one cell population lacks the function of the respective gene and consequently may not contribute to tissue formation, thus leading to developmental abnormalities (Figure 1).

### Table 1: Haploinsufficiency and Disease Association of Human and Mouse Pax Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Human disease</th>
<th>Mouse mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax1</td>
<td>n.d.</td>
<td>Undulated</td>
</tr>
<tr>
<td>Pax2</td>
<td>Renal-coloboma syndrome</td>
<td>Krd</td>
</tr>
<tr>
<td>Pax3</td>
<td>Waardenburg syndrome</td>
<td>Sploch</td>
</tr>
<tr>
<td>Pax6</td>
<td>Aniridia</td>
<td>Peters’ anomaly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Small eye</td>
</tr>
<tr>
<td>Pax8</td>
<td>Hypothyroidism</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

The indicated human disease syndromes and mouse developmental mutants are caused by heterozygous mutations in five different Pax genes (reviewed by Strachan and Read, 1994; Hanson and van Heyningen, 1995; Mansouri et al., 1996). References of the original publications: renal-coloboma syndrome (Sanyanusin et al., 1995), Waardenburg syndrome (Baldwin et al., 1992; Tassabehji et al., 1992), aniridia (Ton et al., 1991), Peters’ anomaly (Hanson et al., 1994), congenital hypothyroidism (Macchia et al., 1998), Undulated (Wilm et al., 1998), Krd (kidney and retinal defects) (Keller et al., 1994), Sploch (Epstein et al., 1991), Small eye (Hill et al., 1991). n.d., not determined.
Hanson and van Heyningen, 1995). Interestingly, the haploinsufficient phenotype of a Pax gene is restricted to only a subset of its expression domains. For instance, heterozygous Pax6 mutations cause malformations only in the eye, although this gene is widely expressed in the developing central nervous system, nose and pancreas in addition to the developing eye (Walther and Gruss, 1991; Calaerts et al., 1997). Hence, the haploinsufficient phenotypes are relatively mild compared to the homozygous null mutations which are lethal for most Pax genes analyzed in the mouse (Mansouri et al., 1996). Moreover, the function of Pax transcription factors appears to be highly dosage-sensitive as indicated by the fact that moderate overexpression of Pax6 generates a similar phenotype as inactivation of one of the two Pax6 alleles (Schedl et al., 1996).

To date, the molecular mechanism responsible for this dosage sensitivity has not yet been elucidated. This phenomenon may be partly explained by the binding site occupancy model (see above; Schedl et al., 1996). However, no evidence for strongly interacting Pax partner proteins exists to date nor is dimerization required for DNA-binding of Pax proteins (Czerny et al., 1993). Hence, it seems unlikely that the function of Pax transcription factors is titrated within the cell by strong protein-protein interactions. Our recent data on the allele-specific regulation of Pax5 (Nutt et al., 1999) raise, however, the possibility that monoallelic expression may be an important mechanism causing the haploinsufficiency of Pax genes.

The Heterozygous Pax5 Mutant Mouse as a Model System to Test the Monoallelic Expression Theory

The Pax5 gene coding for the B cell-specific activator protein (BSAP) is expressed throughout B cell development except in terminally differentiated plasma cells (Adams et al., 1992) (Figure 2A). Targeted inactivation of Pax5 arrests B lymphopoiesis at an early pro-B cell stage in the bone marrow of adult mice, while Pax5 is already required for the generation of the earliest B-lineage committed precursor cells in the fetal liver (Urbánek et al., 1994; Nutt et al., 1997) (see Figure 2A). Pax5-deficient mice also display abnormal patterning of the midbrain and anterior cerebellum (Urbánek et al., 1994) consistent with expression of Pax5 at the midbrain–hindbrain boundary during CNS development (Adams et al., 1992).

The identification of Pax5 target genes has been greatly facilitated by the fact that Pax5-deficient pro-B cells can be readily established as long-term in vitro cultures in the presence of stromal cells and interleukin-7 (Nutt et al., 1997). CD19, coding for a B cell surface protein, is a particularly interesting target gene, as its expression is entirely lost in Pax5-deficient pro-B cells (Nutt et al., 1997). CD19 expression can, however, be fully reactivated in these cells by an estrogen-inducible version (BSAP-ER) of Pax5 (Nutt et al., 1998). Moreover, Pax5 is known to regulate the CD19 gene by binding to a high-affinity site in the -30 promoter region (Kozmik et al., 1992). Hence, CD19 was demonstrated by loss- and gain-of-function experiments to be a bona fide target gene whose expression is strictly dependent on Pax5 function.

A direct test of the monoallelic expression hypothesis depends on the possibility to follow the transcription of individual Pax gene alleles in vivo. Importantly, such a possibility was provided by the analysis of B lymphocytes from heterozygous Pax5 (+/-) mice for the following reasons. First, as stated above, the expression of CD19 is totally dependent on the presence of Pax5, indicating that CD19 is an ideal marker for Pax5 function (Nutt et al., 1997; 1998). Hence, only cells transcribing the wild-type Pax5 allele express CD19 on the surface and can thus be detected as CD19+ cells by anti-CD19 antibody staining and flow cytometric analysis (Figure 2B). Second, the inactivated Pax5 allele contains an in-frame lacZ gene insertion, and hence its transcription leads to the synthesis of β-galactosidase, the activity of which can also be monitored by flow cytometric analysis using a fluorogenic substrate (Figure 2B).

Analysis of the B cell compartments of heterozygous Pax5 (+/-) fetal liver and bone marrow demonstrated that the majority of the pro-B cells express only the CD19 protein, but no β-galactosidase activity (Nutt et al., 1999) (Figure 2D). Hence, most heterozygous pro-B cells express exclusively the wild-type Pax5 allele in vivo, indicating that Pax5 is monoallelically transcribed at the onset of B lymphopoiesis. Monoallelic expression also predominates in mature B cells of the spleen and lymph node, while Pax5 is transcribed from both alleles in pre-B and immature B cells (Nutt et al., 1999) (Figure 2C, D). The pattern of allelic Pax5 transcription is therefore regulated during B cell development. This regulation is furthermore independent of the parental origin of the two Pax5 alleles (Nutt et al., 1999), indicating that the monoallelic expression of Pax5 is not caused by genomic imprinting which normally restricts expression either to the maternal or paternal allele (Barlow, 1995).

The phenomenon of monoallelic Pax5 expression was further investigated by establishing pro-B cell lines from heterozygous mice in vitro. Single-cell cloning experiments and time course analyses demonstrated that individual cell colonies were able to switch expression between alleles within two weeks (Nutt et al., 1999). Moreover, replication timing analyses by fluorescence in situ hybridization (FISH) demonstrated that both Pax5 alleles are synchronously replicated during the S-phase in B cells of heterozygous mice (Nutt et al., 1999). In summary, the allele-specific regulation of Pax5 is stochastic, reversible, independent of parental origin and does not correlate with asynchronous replication in contrast to the monoallelic expression of genomically imprinted genes (Barlow, 1995; Watanabe and Barlow, 1996). As predicted by the model shown in Figure 1, the monoallelic expression of Pax5 generates a haploinsufficient phenotype at the cellular level in heterozygous Pax5 (+/-) mice. Cells which express only the mutant (lacZ) allele are absent in the B-lymphoid lineage past the developmental block of the Pax5 mutation.
Interestingly, the Pax5 (+/−) mouse has been ideal to prove the monoallelic expression theory, although it is phenotypically normal and does not provide a model for a human inherited disease. The main reasons for this apparent discrepancy are the functional redundancy of Pax2 and Pax5 in the developing midbrain (Urbánek et al., 1997; Schwarz et al., 1997) and the fact that a two-fold reduction in B cells is without phenotypic consequence. Hence, the human PAX5 gene could not be linked to a primary immunodeficiency syndrome (Vorechovsky et al., 1995).

(Nutt et al., 1999) (Figure 2C). Interestingly, the Pax5 (+/−) mouse has been ideal to prove the monoallelic expression theory, although it is phenotypically normal and does not provide a model for a human inherited disease. The main reasons for this apparent discrepancy are the functional redundancy of Pax2 and Pax5 in the developing midbrain (Urbánek et al., 1997; Schwarz et al., 1997) and the fact that a two-fold reduction in B cells is without phenotypic consequence. Hence, the human PAX5 gene could not be linked to a primary immunodeficiency syndrome (Vorechovsky et al., 1995).

**Testing the Monoallelic Expression Hypothesis in Genetically Unmanipulated B Cells**

The evidence presented above is persuasive that the two Pax5 alleles are independently regulated in B-lymphoid cells of heterozygous mutant mice. While the B lymphocytes of these mice all transcribe the wild-type Pax5 allele, a variable number of B cells, depending on their developmental stage, also expresses the targeted allele (Figure 2D). These results can be interpreted in two ways. First, Pax5 is regulated in an allele-specific manner during B cell development which is characterized by distinct phases of predominantly mono- or biallelic Pax5 transcription (Nutt et al., 1999). Second, the transcriptional regulation of the
targeted (lacZ) allele has been adversely affected by insertion of the lacZ/neomycin expression cassette. In this case, the mosaic expression pattern of Pax5 may simply reflect dysregulation of the targeted allele.

Evidence against the latter argument is provided by the expression analysis of in vitro cultured Pax5 (–/+ ) pro-B cells (Nutt et al., 1999). We could clearly demonstrate that pro-B cells transcribing the wild-type Pax5 allele in vivo could switch expression from the wild-type to the targeted allele upon in vitro culture, thus indicating that the targeted lacZ allele can be fully active in pro-B cells. Moreover, the wild-type Pax5 allele lacking any genetic modification is transcriptionally silenced under these conditions, while this phenomenon was never observed with cultured pro-B cells from wild-type mice. Nevertheless, the insertion of a neomycin selection cassette upon gene targeting has been shown to affect the transcriptional control of several other genes in the mouse. An illustrative example is the disruption of the myf6 gene, which codes for a bHLH transcription factor involved in myogenesis. Three different myf6 null alleles have been reported which result in distinct phenotypes ranging from complete viability to lethality of the homozygotes (Olson et al., 1996). These phenotypic differences are caused by the specific orientation and position of the inserted neomycin gene which can interfere with cis-acting regulatory elements of the neighbouring myogenic gene myf5 (Olson et al., 1996). Due to the large size of the Pax5 locus (Busslinger et al., 1996), it is unlikely that the insertion of the neomycin gene affects the regulation of an adjacent gene which furthermore would have to be involved in the transcriptional control of Pax5. The neomycin gene insertion could, however, disrupt the regulation of the targeted Pax5 locus itself. Indeed, different locus-specific effects were observed in lymphoid cells, whenever regulatory regions of the immunoglobulin IgH or IgL genes were inactivated either by insertion of a neomycin gene or by ‘clean’ deletion using the Cre-loxP system (Gorman et al., 1996).

To unequivocally prove the monoallelic expression hypothesis, it is therefore important to confirm the data obtained with the heterozygous mouse by analyzing Pax5 expression in genetically unmanipulated B cells containing two equivalent wild-type Pax5 alleles. Two techniques are currently available to detect the expression of wild-type alleles at the single-cell level. Below we will briefly describe the two methods and their results which provided further support for monoallelic expression of Pax5 in B lymphocytes.

... by RNA-FISH

The RNA-FISH method allows the detection of nascent primary transcripts at their site of synthesis in the nucleus. For this purpose, biotinylated DNA probes derived from introns of the gene of interest are hybridized to non-denatured nuclei, and the hybridization signal is subsequently amplified by incubation with fluorescently labelled avidin (Johnson et al., 1991; Panning and Jaenisch, 1996) (Figure 3A). As non-spliced pre-mRNA is primarily found at the site of RNA synthesis in the nucleus, each actively transcribing allele is revealed as a single signal (dot) by the use of intron-specific probes (Figure 3A). One dot per nucleus identifies a monoallelically transcribing cell, while two dots

Fig. 3  Experimental Strategies for Demonstrating Monoallelic Expression in Single Cells.
(A) Schematic diagram of the RNA-FISH method with representative examples of quiescent splenocytes expressing Pax5 mono- or biallelically. (B) Flow diagram of the single-cell RT-PCR technique used to demonstrate monoallelic expression of Pax5 in wild-type B cells of the spleen.
are characteristic of biallelic transcription in a quiescent or G1 cell. This technique has been successfully used to visualize the transcriptional activity of globin genes in erythroid precursor cells (Wijgerde et al., 1995) and to follow the establishment of monoallelic transcription of the Xist gene during X chromosome inactivation (Panning and Jäenisch, 1996). We have recently adapted the RNA-FISH technique to study the transcription pattern of Pax5 in B cells from the spleen of wild-type mice (Nutt et al., 1999).

The Pax5 mRNA is expressed at low abundance in splenic B cells (see below), and hence detection of primary Pax5 transcripts by RNA-FISH is not trivial. Consequently, the RNA-FISH experiments had to be carefully controlled in order to avoid misinterpretation of the data. For instance, inefficient hybridization would inadvertently generate monoallelic expression data, as even potentially biallelic cells may appear monoallelic. To estimate the efficiency of hybridization, we determined the frequency of cells lacking any Pax5 RNA signal, which ideally should be equivalent to the percentage of dead cells measured in the cell population at the onset of RNA analysis. The hybridization efficiency was thus assumed to be close to 100%, if every live cell gave rise to at least one RNA signal. Only the data of RNA-FISH experiments fulfilling this criterion were interpreted. To further assess the hybridization conditions, we performed the same RNA-FISH analysis with a second transcription factor gene, Ikaros, which is expressed in lymphocytes at a similarly low level as Pax5 (Georgopoulos et al., 1992). Interestingly, the Ikaros probe detected 2.4 RNA signals at a high frequency (91%) in the nucleus, indicating that Ikaros is a biallelically transcribed gene. In contrast, monoallelic Pax5 expression was observed in the majority (72%) of the same splenic B cells. A similar proportion (∼75%) of monoallelically transcribing cells was determined by monitoring CD19 and lacZ expression in the heterozygous Pax5 (+/-) spleen which normally contains fewer (−25%) immature B cells (biallelic) and more (∼75%) mature B cells (monoallelic) (Nutt et al., 1999). In conclusion, the RNA-FISH experiments demonstrated that Pax5 is subject to allele-specific regulation in genetically unmanipulated B cells.

... by Single-Cell RT-PCR Analysis

The expression of a given gene can also be studied at the single cell level by a modified two-step RT-PCR protocol (Hu et al., 1997) (Figure 3B). This technique can even be used to discriminate between the expression of individual polymorphic alleles in a single cell. Indeed, single-cell RT-PCR analyses have been successfully utilized to demonstrate monoallelic transcription of the odorant receptor genes in olfactory neurons (Chess et al., 1994) and of the IL-2 and IL-4 genes in T lymphocytes (Holländer et al., 1998; Bix and Locksley, 1998). We also employed this technique to investigate the allele-specific regulation of Pax5 in wild-type B cells. Furthermore, we decided to analyze expression of CD19 as a biallelic control gene in the same RT-PCR experiments. As shown by the RNase protection analysis in Figure 4A, the Pax5 and CD19 genes are expressed at similar levels in pro-B and mature B cells. Quantification of the RNase-protected signals revealed that a single splenic B cell contains as little as 30 and 80 molecules of Pax5 and CD19 mRNA, respectively (Figure 4B; see also legend). Hence it became apparent that the reverse transcriptase reaction had to be quite efficient to ensure the synthesis of several cDNA transcripts starting with so few mRNA templates in a single cell.

To be able to discriminate between transcribed alleles, we have searched for sequence polymorphisms in the CD19 and Pax5 mRNA of two related subspecies of the mouse, i.e., in the laboratory strain C3H and in the Mus musculus strain PWD (Forejt and Gregorová, 1992). cDNA cloning and sequence analysis revealed a cluster of changes in the 3’ non-coding region of Pax5, which were used to generate PWD- and C3H-specific oligonucleotide probes (Figure 5A). cDNA products amplified from B cells of the PWD and C3H strains could be readily distinguished by stringent hybridization with these probes (Figure 5B). The coding region of the CD19 gene was furthermore shown to differ at position 665 by a C-to-T transition, which generated a BgIII restriction site in the PWD sequence (Figure 5C). Hence, BgIII digestion could be conveniently used to discriminate between the two allelic CD19 transcripts that were amplified from B cells of F1 (C3H × PWD) hybrid mice (Figure 5D).

The two-step PCR protocol, which relies on the use of nested oligonucleotide primers (Figure 3B), is extremely sensitive, as it can detect a single cDNA molecule present in the initial RT reaction. Hence, a critical parameter of the single-cell RT-PCR experiment is the number of cDNA molecules present in a single B lymphocyte. (A) RNase protection analysis. CD19 and Pax5 mRNA levels were quantitated by RNase protection assay in total RNA (10 μg) prepared from B220+ sorted splenocytes and in vitro cultured pro-B cells. The size of the RNase-protected signal is indicated in nucleotides to the right. (B) Quantitation of the RNase protection data. The number of mRNA molecules per B lymphocyte was calculated from the known number of B cells required to obtain 10 μg of total RNA, the specific activity of each riboprobe and the radioactivity present in the RNase-protected signals (measured by phosphorimager analysis).

---

Fig. 4 Quantitation of the Pax5 and CD19 mRNA Molecules Present in a Single B Lymphocyte.

(A) RNase protection analysis. CD19 and Pax5 mRNA levels were quantitated by RNase protection assay in total RNA (10 μg) prepared from B220+ sorted splenocytes and in vitro cultured pro-B cells. The size of the RNase-protected signal is indicated in nucleotides to the right. (B) Quantitation of the RNase protection data. The number of mRNA molecules per B lymphocyte was calculated from the known number of B cells required to obtain 10 μg of total RNA, the specific activity of each riboprobe and the radioactivity present in the RNase-protected signals (measured by phosphorimager analysis).
transcripts which are synthesized by reverse transcriptase prior to exponential amplification. Inefficient cDNA synthesis may generate only one cDNA transcript even from biallelically transcribing cells, thus creating an artifact, as every gene examined would appear to be monoallelically expressed. Using CD19 cDNA as a template, we optimized several parameters of the single-cell RT-PCR assay to achieve the desired goal of amplifying both allelic transcripts from a single B cell of an F1 hybrid mouse (for details see legend to Figure 6). In particular, two modifications proved to be crucial, as they increased the number of amplification-positive cells up to 95% and allowed the detection of both allelic CD19 transcripts in > 90% of these cells (Figure 6A). First, the use of gene-specific oligonucleotides as opposed to random hexamer or oligo(dT) primers led to a significant increase in the number of cells from which CD19 transcripts could be amplified. Second, RNase H digestion of the RNA-DNA hybrids of the RT reaction considerably improved the efficiency of PCR amplification. Importantly, the two CD19 transcripts were amplified in an unbiased manner, as both alleles were detected with equal efficiency upon serial dilution of the corresponding cDNA mixture (Figure 5E).

The result of a typical single-cell RT-PCR experiment, which was performed under optimized conditions with individual B220+ IgM+ splenocytes, is shown in Figure 6. Both allelic CD19 transcripts were detected in the majority of cells, as expected for a biallelically expressed gene. However, only 38% of the CD19 biallelically transcribing cells expressed both Pax5 alleles (Figure 6B,C). Hence, the single-cell RT-PCR method revealed monoallelic Pax5 expression in 62% of all splenic B cells which is in good agreement with the RNA-FISH data (72%) and flow cytometric analysis of the Pax5 (+/-) mice (~75%).

Methods: F1 hybrid off-spring was obtained by intercrossing mice of the inbred laboratory strain C3H. The following oligonucleotides were used for single-cell RT-PCR analysis. mouse Pax5:

Methods: F1 hybrid off-spring was obtained by intercrossing mice of the inbred strain PWD (Forej and Gregorová, 1992) and the inbred laboratory strain C3H. The following oligonucleotides were used for single-cell RT-PCR analysis. mouse Pax5: (a) 5’-CTAAGCACCACAGACACACC-3’; (b) 5’-GAGCTTGCGAGCACCAACCC-3’; (c) 5’-CTCCTGCTGCTGAGCAGGGT-3’; (d) 5’-TCAAGGGTTGAGGAGCTGCC-3’; mouse CD19: (a) 5’-GAATCGTCTGAGGAGCAGGGT-3’; (b) 5’-CCACAGGCTTCACCTCTGTC-3’; (c) 5’-GACCTGACCTGGGACTATCC-3’; (d) 5’-CTCAACAGGCAAGCCACACTG-3’. Discrimination of polymorphic Pax5 and CD19 transcripts: the allelic origin of the CD19 PCR product was determined by BglII digestion followed by gel electrophoretic analysis. Approximately 100 ng of the Pax5 PCR product was denatured by dilution in 0.4 M NaOH and heating to 100°C for 10 min. The denatured DNA fragments were then slot-blotted onto nylon membranes and hybridized to end-labeled allele-specific or common oligonucleotide probes shown by brackets in (A). The specificity of allele detection was achieved by overnight hybridization in 10% formamide / 6 M urea at 60°C and exposure to X-ray film for 1 h at –70°C.

The sizes of the PCR product and its BglII restriction fragments are indicated in base pairs (bp). (B) BglII digestion of CD19 PCR fragments amplified from the spleen of PWD, C3H and F1 (PWD x C3H) hybrid mice. (E) Unbiased PCR amplification of the polymorphic CD19 transcripts. CD19 cDNA (100 ng) derived from the PWD and C3H parental strains was mixed, serially diluted and then subjected to the 2-step PCR regimen described in the legend to Figure 6. PCR products were digested with BglII and analyzed on a 1.5% agarose gel.

Fig. 5 Identification of Nucleotide Polymorphisms in Pax5 and CD19 mRNAs.

(A) Nucleotide polymorphisms were identified in the 3’ non-coding sequence of Pax5 exon 10 by RT-PCR amplification and sequencing of a 332-bp cDNA fragment from the spleen of C3H and PWD mice. Nucleotide positions correspond to the published mouse Pax5 cDNA sequence (Adams et al., 1992), and arrowheads denote the positions of introns. (B) The oligonucleotide probes shown by brackets in (A) were used for high-stringency slot-blot hybridization of Pax5 PCR products which were amplified from B cells of PWD and C3H mice. (C) A polymorphic BglII restriction site in the coding region of CD19 (Zhou et al., 1991) facilitates rapid determination of the allelic origin of CD19 transcripts by BglII digestion. (D) BglII digestion of CD19 PCR fragments amplified from the spleen of PWD, C3H and F1 (PWD x C3H) hybrid mice. (E) Unbiased PCR amplification of the polymorphic CD19 transcripts. CD19 cDNA (100 ng) derived from the PWD and C3H parental strains was mixed, serially diluted and then subjected to the 2-step PCR regimen described in the legend to Figure 6. PCR products were digested with BglII and analyzed on a 1.5% agarose gel.
Monoallelic Expression as a Result of Stochastic Activation of Pax5 at Lineage Commitment

Pax5 is the first transcription factor gene which has been shown to be predominantly expressed from only one of its two alleles in a stochastic and reversible manner. This allele-specific regulation generates a haploinsufficient phenotype at the cellular level in B-lymphoid tissues of Pax5 (+/−) mice (Nutt et al., 1999). However, for reasons discussed above, heterozygous mutation of the Pax5 gene has no phenotypic consequences (Urbánek et al., 1994) and hence does not provide a model for a human disease syndrome (Vorechovsky et al., 1995). In contrast, heterozygous mutations of the closely related PAX2 and PAX8 genes of the same Pax subfamily are known to cause the renal-coloboma syndrome (Sanyanusin et al., 1995) and congenital hypothyroidism (Macchia et al., 1998), respectively. It is therefore tempting to speculate that allele-specific transcriptional regulation is responsible for the susceptibility of these and other mammalian Pax genes to single mutations. A model of stochastic allele-specific regulation of Pax genes could account for two aspects of Pax-related disease syndromes; i.e. their phenotypic variability and tissue preference. The severity of a heterozygous Pax mutation would be expected to depend on the stochastically determined ratio of cells activating either the wild-type or mutant allele. An increased frequency (> 50%) of cells expressing the wild-type allele may result in a milder phenotype, while a higher proportion (> 50%) of cells activating the mutant allele should cause a more severe phenotype. Moreover, the independent regulation of the two Pax5 alleles during B cell development (Figure 2) suggests that other Pax genes may also be transcribed monoallelically in some tissues and biallelically in others, thus explaining why haploinsufficient phenotypes are observed only in a subset of the Pax expression domains.

### Table: Summary of Pax5 transcription in individual B cells which were all shown to be biallelically or monoallelically expressing CD19.

<table>
<thead>
<tr>
<th>mCD19</th>
<th>mPax5</th>
<th>C3H</th>
<th>PWD</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>M</td>
<td>B</td>
<td>M</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>M</td>
<td>B</td>
</tr>
<tr>
<td>M</td>
<td>M</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>M</td>
<td>B</td>
<td>M</td>
<td>M</td>
</tr>
<tr>
<td>M</td>
<td>M</td>
<td>C3H</td>
<td>PWD</td>
</tr>
</tbody>
</table>

(Nutt et al., 1999). In summary, the results obtained with these fundamentally different methods support the same conclusion that Pax5 is subject to allele-specific regulation in B lymphocytes.

Methods: Single-cell RT-PCR analysis: Single-cell suspensions prepared from the spleen of F1 (PWD × C3H) hybrid mice were incubated with 0.5 U RNase H (Boehringer Mannheim) for 30 minutes at 37°C and again heating to 70°C for 10 minutes. 76 μl of the PCR mixture (1 × PCR buffer (Boehringer Mannheim) containing 0.8 mM dNTPs, 0.75 μg of the Pax5 and CD19 PCR primers α and γ, and 4 U Taq DNA polymerase (Boehringer Mannheim)) were then added to the RNase H-treated RT reaction followed by 25 cycles of PCR amplification. 2 μl of the first PCR reaction was then used for separate amplification of Pax5 and CD19 cDNA by utilizing the respective nested primer combinations β and δ for additional 25 cycles. The 332-bp Pax5 and 423-bp CD19 cDNA fragments could be readily visualized by analyzing 20 μl of the second PCR reaction on a 1.5% agarose gel.

---

### Fig. 6: Single-Cell RT-PCR Analysis of Pax5 and CD19 Expression in Splenic B Cells.

(A) Biallelic transcription of the polymorphic CD19 alleles in individual B cells of F1 (PWD × C3H) hybrid mice. The PCR products of PWD and C3H origin were discriminated by BglII digestion. PCR reactions containing only one of the two allelic CD19 transcripts (denoted by asterisks) were excluded from further analysis of Pax5 expression. (B) Slot-blot analysis of the RT-PCR reactions of 32 individual B cells. The PCR reactions of only those cells which were determined to be biallelic for CD19 transcription were analyzed by slot-blot hybridization with PWD- and C3H-specific probes (see legend to Figure 5). Pax5 cDNA derived from the spleen of parental PWD and C3H mice was used as specificity control. M, monoallelic; B, biallelic. (C) Summary of Pax5 transcription in 75 individual B cells which were all shown to be biallelically expressing CD19.

<table>
<thead>
<tr>
<th>monoallelic (M)</th>
<th>biallelic (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>46 (62%)</td>
<td>29 (38%)</td>
</tr>
</tbody>
</table>

---
Monoallelic Expression May be More Common in Mammals than Previously Anticipated

The majority of mutations in mouse and human genes are recessive, indicating that expression from both alleles is not essential for the function of most genes in vivo. A small proportion of mammalian genes (0.1% - 0.2%) is, however, monoallelically expressed as a result of genomic imprinting, while the extent of non-imprinted genes displaying monoallelic expression is not yet known (Watanabe and Barlow, 1996). Random monoallelic expression has been first described for the family of odorant receptor genes (Chess et al., 1994). In this case, monoallelic transcription is a mechanism to ensure that each individual sensory neuron of the olfactory system expresses only a single receptor on the surface. The lymphoid system, which is more amenable to analysis at the single cell level, has provided further insight into the regulation of monoallelic gene expression. The Ly49 gene family codes for MHC class I-specific inhibitory receptors which are expressed on the surface of natural killer (NK) cells where they participate in the discrimination of self from nonself in innate immunity (Long and Wagtman, 1997). Interestingly, individual alleles of these Ly49 genes are stochastically chosen for stable expression in different NK cells. Monoallelic expression of the Ly49 genes is therefore thought to be essential for generating a diverse repertoire of NK cell receptors (Held et al., 1995; Held and Raulet, 1997; Held and Kunz, 1998). More recently, the analysis of wild-type T cells and heterozygous mutant mice has revealed that individual alleles of the IL-2 and IL-4 genes are independently regulated by a stochastic process in T helper cells (Hollandner et al., 1998; Bix and Locksley, 1998; Rivière et al., 1998). Allele-specific regulation may therefore establish a diverse repertoire of combinatorially assorted cytokine expression patterns in T lymphocytes (Bix and Locksley, 1998; Rivière et al., 1998).

Several human disease syndromes are caused by heterozygous mutations in transcription factor genes other than PAX genes (MITF, HOXD13, EMX2, GLI3, THRB, HMGIC, SOX9, SOX10, CBP, EYA1, CBFA1) (Engelkamp and van Heyningen, 1996, and references therein). It will be important to see whether monoallelic expression of some of these genes may be the underlying disease-causing mechanism. An interesting case in this respect is CBFA1, which upon mutation causes the cleidocranial dysplasia syndrome (Lee et al., 1997; Mundlos et al., 1997). Targeted gene inactivation in the mouse identified Cbfa1 as a key regulator of osteoblast differentiation (Otto et al., 1997; Komori et al., 1997), suggesting that this transcription factor may be essential for commitment to the osteoblast lineage similar to the role of Pax5 in B cell development.

Acknowledgements

This review is a late contribution to the highlight issue of March ’99 which is dedicated to Max L. Birnstiel, the founding director of the Research Institute of Molecular Pathology in Vienna. Meinrad Busslinger would like to thank Max L. Birnstiel for many years of scientific advice, support and encouragement. We are furthermore grateful to E. ten Boekel for advice on single-cell RT-PCR, to A. Souabni for cloning polymorphic Pax5 and CD19 cDNA, to P. Steinlein for FACS sorting and to L. Huber for critical reading of the manuscript. This work was supported by the Research Insti-
tute of Molecular Pathology and by a grant from the Austrian Industrial Research Promotion fund.

References


