# *CHAPTER 2*
# *STOCHASTIC PATTERNS IN GLOBIN GENE EXPRESSION ARE ESTABLISHED PRIOR TO TRANSCRIPTIONAL ACTIVATION AND ARE CLONALLY INHERITED*

Mariken de Krom, Mariette van de Corput, Marieke von Lindern, Frank Grosveld and John Strouboulis

*Molecular Cell* Vol **9**, 1319-1326 (2002)

## Summary

We have undertaken a detailed characterization of mouse globin gene expression patterns in the nucleus and cytoplasm of single erythroid cells. We demonstrate an imbalance of α- versus β-globin expression in a significant proportion of cells both in nuclear transcription patterns and cytoplasmic mRNA levels. Clonal cell analysis showed these expression patterns to be clonally inherited, while analysis of a multicopy transgenic locus showed an all-or-none effect in the activation of all the genes in one locus. These data provide strong evidence for a stochastic basis of globin gene activation resulting in heritable all-or-none expression patterns.

## Introduction

A stochastic basis for transcriptional activation can be viewed as a sequence of events that combines a random component, such as transcription factor binding, with a selective step, such as cell commitment, so that only certain outcomes of the random event persist (Bateson, 1979). In this model, cis regulatory elements, such as enhancers, or an increase in an inducing signal, such as a transcription factor, increase the probability that a gene will be transcriptionally activated (Ko, 1990).

One implication of the probabilistic nature of a stochastic model is that the individual alleles of a gene will be regulated independently. Support for this comes from random X chromosome inactivation in dosage compensation (reviewed in Goto and Monk, 1998), from allelic exclusion of immunoglobulin genes in immune responses (reviewed in Nemazee, 2000), and from random monoallelic expression of genes involved in lineage commitment, such as olfactory receptors or T and NK cell receptors (Chess *et al*., 1994; Held *et al*., 1999), and in cellular responses to external stimuli, such as the cytokine genes IL-2 and IL-4 in activated T cells (Hollander *et al*., 1998; Bix and Locksley, 1998; Riviere *et al*., 1998). In all these cases, stochastic choices in allelic expression patterns of a limited number of genes give rise to a diverse repertoire of cells, each with a highly restricted specificity. Upon stimulation/induction, the cell is committed and clonally expanded to provide a specific response (Bix and Locksley, 1998; Chess, 1998; Shulman and Wu, 1999; Held *et al*., 1999). The advantage of a stochastic mechanism is that it creates diversity by allowing choice at low cost. While this predicts that individual cells will vary, the net result for a cell population will be a stable outcome since large numbers of cells, each with the same probabilities for a particular event, are involved (Michaelson, 1993).

The nuclear expression patterns of α- and β-globin genes appear consistent with the predictions of a stochastic model (Wijgerde *et al*., 1995, 1996; Gribnau *et al*., 1998; Trimborn *et al*., 1999). Total mouse globin mRNA levels reflected the number of actively transcribing genes in nuclei, suggesting that, when transcribing, genes are fully on. Moreover, different combinations of mouse globin expression patterns were observed, suggesting that each globin locus is independently regulated.

Here, we examine the apparent stochastic basis for globin gene activation by calculating the proportion of genetically identical mouse erythroid cells that fail to activate one or both of the α- and β-globin alleles. This results in an imbalance in α- versus β-globin transcription patterns which is reflected in α/β mRNA ratios in the cytoplasm. We also provide evidence that globin allelic transcription patterns are clonally inherited and that decisions for transcription occur in an all-or-none fashion at a step prior to transcriptional activation.

## Results

### *Independent Activation of Globin Gene Alleles*

*In situ* hybridisation was used to detect globin nuclear primary transcripts and/or cytoplasmic mRNA in 14.5 days post coitus (dpc) mouse foetal liver cells, which express the adult α-globin and $\beta_{maj}$- and $\beta_{min}$-globin genes (Trimborn *et al*., 1999). The efficiency of hybridisation in these assays was

shown to be ~99% (data not shown). We then measured α- and β-globin primary transcription foci using intron-specific probes. A range of allelic expression combinations was observed with the majority of cells expressing two α and two β alleles (Figures 1A–1B). Active alleles per cell were counted and grouped as shown in the table in Figure 1. From this, two important points emerge. First, while >73% of cells (Figure 1C, column 4) exhibit balanced transcription of 2α- and 2β-globin alleles, the remainder of the cells (~17% of the total, Figure 1C) fail to transcribe at least one globin allele. Second, α- and β-globin transcription is clearly unequal, as more cells express predominantly or exclusively α-globin, compared to cells expressing mostly β-globin (18% versus 5%; Figure 1C, columns 1 + 2 + 3 versus columns 6 + 7 + 8). This is contrary to predictions if one assumed an equal activation potential for α- and β-globin alleles. The number of cells expressing only β alleles is not statistically significant, as it is below the 1% limit of hybridization efficiency.

We next calculated the probability of activation for each α (pα)- and β (pβ)-globin gene. For α-globin, pα can be calculated from the 2α-expressing cells ($p\alpha^2 = 87.2\%$ [sum of columns 2–4, Figure 1C]; therefore pα is 93.4%). Similarly, pβ is calculated to be 88.4% (sum of columns 4, 6, and 7, Figure 1C). This would result in an expected 20.5% or 1.3% of cells expressing a single or no β gene, respectively. In fact, the observed number of single or no β allele-expressing cells is 12.4% and 9.3%, respectively (columns 3, 5, and 8, and 1 and 2, respectively, Figure 1C). This discrepancy is caused by the high incidence of α-only-expressing cells. The trivial explanation for the higher number of α-only-expressing cells (contamination with embryonic cells expressing α-globin) was excluded by *in situ* hybridisation of 14.5dpc foetal liver cells using probes specific for mouse α-globin and the mouse εy-embryonic-globin gene (data not shown).

### *α-Globin Is Activated Earlier Than β- in Erythroid Differentiation*

We can think of three plausible explanations for the high number of α-only-expressing cells. First, mouse α-globin genes could, on average, be transcribed earlier than β genes in erythroid cell maturation. Second, the α genes could be expressed at a specific cell cycle stage when the β genes are not active (e.g., in G1 phase). Third, the α-only cells could result from discontinuous transcription, as has been previously described for the α- and β-globin genes (Wijgerde *et al*., 1995; Gribnau *et al*., 1998), with the β genes transcribing less frequently than the α genes. This would result in cells with α on in the nucleus, but with the β gene(s) temporarily off.

To address the first possibility, we size fractionated mouse foetal liver cells into large immature erythroid cells and small, more mature erythroid cells. *In situ* hybridisation with α- and β-globin exon probes, which detect cytoplasmic mRNA and nuclear primary transcripts, showed that the large immature erythroid cells have one or two α-globin primary transcripts and little or no cytoplasmic α mRNA. No β primary transcripts or mRNA were present (Figure 2A). Thus, these cells have just started transcribing the α-globin genes ahead of the β genes and have not yet built up detectable mRNA in the cytoplasm. The vast majority of the more mature cells contain both α- and β-globin primary transcripts and mRNA (Figure 2B). Re-counting shows a similar distribution of primary transcripts as in Figure 1C, with the exception that the α-only cells or cells expressing more α than β have now decreased (Figure 1D). We conclude that the α-globin locus becomes transcriptionally active before the β-globin locus.

The observation that cells with α-only signals in the nucleus had little or no α signal and no β signal in the cytoplasm argues against the possibilities that the α genes are activated at a stage of the cell cycle when β is off or that the α genes are transcribed more frequently than β. If these possibilities were true, these cells would have both α and β cytoplasmic mRNA synthesized in previous cell cycles. Thus, the α-only cells are primarily due to the earlier activation of β-globin during erythroid cell differentiation.

We recalculated the activation probabilities for both α- and β-globin genes, using the mature erythroid cell fraction. We now found pα to be 94% ($p\alpha^2 = 88.4$, columns 2–4, Figure 1D). Conversely, pβ in fractionated cells is 96% ($p\beta^2 = 92.2$, columns 4, 6, and 7, Figure 1D). This predicts 7.7% of cells expressing a single β allele, which agrees with the 7.6% observed (columns 3, 5, and 8, Figure 1D). The non-β-expressing cells are predicted to be 0.16%, with 0.21% observed (columns 1,

Figure 1D). Hence, we conclude that after removal of the early erythroid cells, α and β expression patterns closely resemble a stochastic distribution, with a very similar probability for α- and β-globin activation.

### *Globin Primary Transcript Imbalance Is Maintained in the Cytoplasm*

The imbalance in α versus β allelic expression in the nucleus of erythroid cells is consistent with a stochastic basis in gene activation. However, it cannot be excluded that the distribution of nuclear transcription patterns observed may be due to discontinuous transcription where an α or a β gene may be temporarily off. If the distribution of nuclear primary transcripts is reflected in α/β mRNA ratios in the cytoplasm, it would confirm that the nuclear transcription patterns are stable and become fixed and would strongly argue for a stochastic activation mechanism.

We carried out *in situ* hybridisation on 14.5dpc foetal liver cells using exon-specific probes and quantitated the relative abundance of cytoplasmic α- and β-globin mRNA (Figures 2C–2F). The distribution obtained for α/β mRNA ratios is shown in Figure 1E. We found 26.7% of the cells having an α/β imbalance. Of those, 15.7% have α > β, with 5.2% of the cells having α/β > 2. The latter group very likely represents the cells that in the primary transcript analysis were expressing only α alleles or 2α1β alleles.

The overall distribution of α/β-globin mRNA ratios agrees with that obtained for the primary transcript analysis, except that the number of cells with α-only primary transcripts is higher than the number of cells with high α/β mRNA ratio in the cytoplasm, because this group also contains the early α-only-expressing cells which have undetectable levels of cytoplasmic α-globin mRNA. Thus, the distribution of globin mRNA in the cytoplasm reflects that of primary transcripts in the nucleus. We conclude that the primary transcript patterns observed in the nucleus cannot be explained by genes being temporarily off but instead demonstrate a stochastic basis for α- and β-globin gene activation.

### *Cellular Globin Expression Patterns Become Fixed during Differentiation*

The agreement of nuclear transcription patterns with mRNA levels suggests that early stochastic choices become fixed in differentiated erythroid cells. We therefore carried out methylcellulose colony assays on single immature erythroid cells, purified under conditions of cell differentiation that allow the visualization of globin transcription before cell enucleation. Individual colonies were harvested after around five divisions (32 cells) and assayed for globin primary transcript expression patterns. If choices in globin expression are not fixed and remain purely stochastic during differentiation, one would expect each colony to display the distribution of nuclear expression patterns observed in the bulk of foetal liver cells (Figure 1C). If, instead, stochastic decisions become fixed, a deviation from the expected stochastic patterns would be predicted (Figure 3A).

We find that the distribution of expression patterns of all the cells assayed in all the colonies is no different to that observed with bulk foetal liver cells. A stochastic model based on the calculated probabilities for α or β gene activation in late erythroid cells (94% and 96%, respectively) would predict the following: in 3.5% of the clones, 100% of the cells express 2α2β; in 21.2% of the clones, 90%of the cells express 2α2β (with the remaining having a different pattern); in 61.6% of the clones, 80% of the cells express 2α2β; in 11.9% of the clones, 70% of the cells express 2α2β, while only 0.6% of the colonies would have cell populations in which less than 60% of the cells express 2α2β. This would lead to a normal distribution (Figure 3B, black bars and curve). Instead, we find a totally different distribution (Figure 3B, gray bars). Many more clones than expected exclusively contain only 2α2β cells or have less than 60% of the cells per clone expressing 2α2β. Many fewer than expected clones contain 80% of cells expressing 2α2β (Figure 3B, gray bars). Importantly, all cells in a colony with patterns other than 2α2β had the same pattern, with two exceptions that contained two patterns different from 2α2β. This can be explained if two or three choices were made in these colonies and subsequently fixed (Figure 3A, right panel).

*All-or-None Activation of Human β-Globin Locus Transgenes*

In the experiments above, we measured transcriptional activation as the read-out for the whole process. We therefore asked whether the stochastic distribution we observe in globin expression patterns is determined prior to or at the transcriptional activation step by examining the activation of three copies of a 180 kb human β-globin locus PAC integrated in a head-to-tail fashion in transgenic mouse lines. If stochastic expression is determined at the transcriptional activation step, each β-locus in the three linked loci would transcribe independently, and, as a result, a substantial number of cells would express only one or two of the three β-loci.

To test this, we used two different three-copy transgenic mouse lines, one bearing a telomeric integration while the other line carried a nontelomeric, noncentromeric insertion. We assumed a similar probability of activation for human β-globin and mouse β-globin (pβ = 88% in nonfractionated cells), as they have equal expression levels per gene copy. If stochastic distribution is determined at the step of transcriptional activation, we would predict a 68% chance for all three loci to be active in the nucleus of one cell ($p\beta^3 = [0.88]^3 = 0.68$). The probability for all three loci to be active in, e.g., ten cells, is around 2% ($[0.68]^{10} = 0.02$).

Foetal liver cells from the three-copy transgenic lines were probed by *in situ* hybridisation with either a 12 kb LCR-derived probe (Figure 4A) or the entire 180 kb PACderived human β-globin locus (Figure 4E1). The LCR probe clearly shows the three separate LCR signals 180 kb apart. Hybridization with the PAC insert shows a brighter, more contiguous signal as more DNA is hybridized. We did a primary transcript hybridization to detect mouse α- and human β-locus transcription. Three types of cells were observed (Figure 4B): some had no α or β signals and are presumably nonerythroid or early erythroid cells (white arrow head, Figure 4B); 19% (telomeric line) and 17% (nontelomeric line) of the remaining cells had only α signals (Figure 4B, solid white arrow); while 81% and 83% of the remaining cells showed α and human β signals (Figure 4B, open white arrow). This fits well with the number predicted from the table in Figure 1, considering that these mice are heterozygous for the human β-globin locus. Deconvolution of confocal images showed the human primary transcripts appearing as a triple signal (Figure 4B, CCD image open arrow; Figure 4C, confocal image, two focal planes merged), while rotation of the nucleus showed the triple signals as short tracts of primary transcript (Figure 4D, confocal image, one focal plane).

We next analyzed 52 human β-expressing cells from the nontelomeric integrant by confocal microscopy for RNA (γ and β intron probes) and DNA (entire PAC). In 49 cells, primary transcripts were visualized as three discernible RNA "tracks" coincident with the human β locus DNA (Figures 4E1–4E3). We did not observe any cell with a single primary transcript track. The remaining 3 (out of 52) cells had coincident primary transcript and DNA signals which, under the confocal microscope, appeared as single large bright foci rather than tracks, indicating a less extended conformation of the DNA. In 9 out of 12 nonexpressing cells, we also observed three DNA signals, indicative of an extended conformation. Essentially identical results were obtained with the telomeric integrant line (n = 30 cells).

In summary, when human β-globin is transcribed, all three loci are active in an all-or-none fashion. We conclude that the stochastic distribution we observe for the human β-globin locus is determined at a step prior to actual transcriptional activation.

## Discussion

Our analysis revealed that a significant proportion of erythroid cells (23.2%) show an imbalance of 2α-versus 2β-globin gene expression, which is maintained in the levels of α- and β-globin cytoplasmic mRNA. In addition, patterns of expression were clonally inherited during erythroid cell differentiation, suggesting that when an allele is off, it remains off. The decision for expression appears to occur in an all-or-none fashion at a step prior to transcriptional activation. These observations have implications for globin gene regulation and the basis of gene activation in general.

### *A Stochastic Basis for Globin Gene Activation*

The α- and β-globin gene loci reside on different chromosomes and are regulated differently and independently. The α-globin locus, unlike the β-locus, lies in an early replicating, constitutively open chromatin domain and does not appear to require the presence in cis of a complete locus control region (LCR), an element required for the continued activation of the β-globin locus (review Higgs *et al*., 1998; Grosveld, 1999). This suggests that the α-globin locus may require fewer steps for activation from its silent state. In addition, there are differences in the transcription factor requirements of the two geneclusters.

The data on globin gene expression patterns presented here provide strong evidence for a stochastic model for transcriptional activation of both globin geneclusters. The higher probability for α-globin activation in the total cell population could result from a lower threshold for productive transcriptional activation, for example, by having more transcription factor binding sites. The constitutively open chromatin configuration of the α-locus could also place it in a nuclear subdomain (Brown *et al*., 2001) that may have a higher diffusion rate for transcription factors (Misteli, 2001). Consequently, α genes would, on average, be transcribed earlier than β genes during differentiation, while at later stages a much smaller proportion of α-only cells will remain. This was observed when we fractionated erythroid cells.

These observations could also be explained by a graded model, in which two alleles of a gene are equally regulated in a graded manner in response to an inducer. However, a graded model would need to invoke additional parameters to account for the lack of transcribing alleles in a significant proportion of the cells. Even so, this model would predict no globin mRNA imbalance in the cytoplasm, as opposed to what is observed in the nucleus. The agreement between primary transcript patterns and cytoplasmic mRNA levels, therefore, argues against the graded response model in this system.

### *Stochastic Expression Choices Are Fixed prior to Transcriptional Activation of a Locus*

We found that mouse globin allelic expression patterns are clonally inherited in differentiated cells. We also showed that, at least for a human multicopy β-globin transgenic locus, the decision to express occurs as an "all-or-none" choice prior to transcriptional activation, probably on a basis of mass action (Grosveld, 1999). Heritable on-or-off decisions are also observed in position effect variegation (PEV) and are thought to be regulated at the level of chromatin structure and subnuclear localization (Wakimoto, 1998; Milot *et al*., 1996). In PEV, mass action has been invoked to account for the stochastic, long-range spreading of heterochromatin in an all-or-none event, which may close down the expression of adjacent genes (Locke *et al*., 1988). Our results suggest that mass action in the β-globin locus works in the opposite direction, in that it opens up chromatin in the locus prior to transcriptional activation. Genetic evidence supports the β-locus LCR as playing a vital role in this all-or-none effect (Grosveld, 1999). To complete the parallel between PEV and globin locus activation, one would need to determine whether the chromatin of nonexpressing globin loci have failed to open, as has been observed in PEV (Wallrath and Elgin, 1995). However, this is technically unfeasible in the current *in vivo* system. Nevertheless, we suggest that the initial choice of activating or not a globin locus is made in an early window of opportunity and subsequently fixed by an as yet unknown parameter at the chromatin level.

### *Globin Gene Expression and Other Stochastic Activation Phenomena*

Stochastic choices in allelic expression have been described in X-inactivation, allelic exclusion, and monoallelic expression and have been proposed to play a role in giving rise to a diverse repertoire of cells, each with a highly restricted specificity to provide a specific response to a stimulus. Globin gene expression, however, does not fit the pattern set by the examples of stochastic choices above. Globin genes are transcribed at high levels in the erythroid lineage with the aim of synthesizing α- and β-globin protein chains in a 1:1 ratio. These cells are fully committed to haemoglobin accumulation and enucleation to form mature erythrocytes; hence they are not part of a cellular response pathway and play no role in cell lineage commitment decisions. It is in fact a disadvantage to the cell to not express

balanced α- and β-globin chains, as it shortens its lifetime. The α- and β-loci have therefore evolved a system whereby the probability of activation is sufficiently high to give a 1:1 ratio of α and β product in the total cell population at a loss of some non-1:1 cells that is sufficiently small so as not to result in a selective disadvantage.

It is tempting to suggest that genes in other tissues may not express for stochastic reasons, even if this were to occur less frequently than observed for the globin genes. For example, even if the chance for activating a gene were 99.9%, one in a million cells would express neither allele of that gene. Such phenomena would be unnoticed if small numbers of cells were assayed, but could play a role when large numbers of cells in a tissue are involved. For example, the failure to activate a tumour suppressor gene in the skin (even in one in a million cells) would increase the chance for transformation in a substantial number of cells. For that reason (among others), additional cellular mechanisms may have evolved to control the phenotypic consequences of such events.

## Material and Methods

### *In Situ Hybridisation*

Mouse 14.5dpc foetal liver cells were isolated, fixed, hybridised, and detected as previously described (Wijgerde *et al*., 1995; van de Corput and Grosveld, 2001). Probe sequences and hybridisation efficiencies have been previously reported (Gribnau *et al*., 1998; Trimborn *et al*., 1999). As control for colour and labelling differences, identical *in situ* hybridisations were performed reversing the fluorescent labels for the α- and β- globin intron and exon probes, yielding similar results. To compare levels of two mRNAs in single cells, a control two-colour mRNA *in situ* hybridisation was done using a single probe mixture with half of the probe mix labelled with digoxigenin (DIG) and the other half with dinitrophenol (DNP). Measuring red and green signal intensities in the cytoplasm revealed the background signal and the distribution intrinsic to the 1:1 ratio of red/ green signal for the same mRNA (Figures 2C and 2F) serving as the baseline distribution for equal amounts of two different mRNAs. Fluorescence was detected by epifluorescence/CCD or confocal microscopy.

DNA-RNA *in situ* hybridisation of the three-copy human β-globin locus was essentially as described by van Raamsdonk and Tilghman (2001). Human γ- and β-globin intron-specific biotinylated probes were used together with mouse α-globin DIG-labeled intron probes. DNA FISH probes included DIG-labelled human β-globin locus PAC insert and a 12 kb LCR probe fragment. Cells were analysed by epifluorescence microscopy or/and confocal laser scanning microscopy, imaging software provided by the Zeiss LSM 510, and deconvolution software (Scientific Volume Imaging, Huygens v. 2.0).

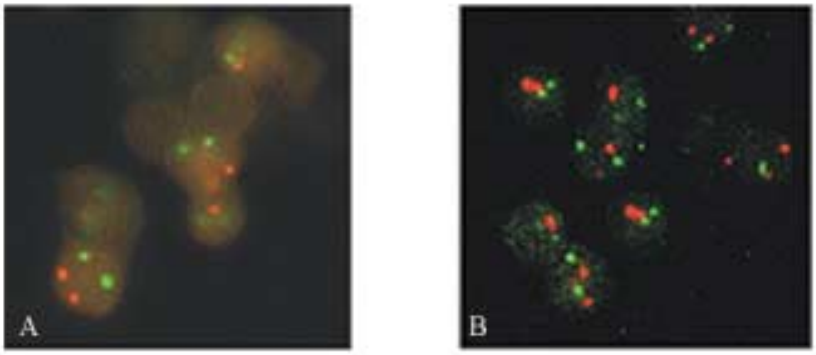
### *Percoll Density Gradient*

Mouse 14.5dpc foetal liver cell suspensions were layered on top of a 2 ml cushion of Percoll (SIGMA-Aldrich) with a density of 1.0735 g/ml and centrifuged at room temperature, 10 min at 2000 rpm (no break for last 300 rpm). The top and bottom fraction contained the immature (larger) and mature (smaller) erythroid cells, respectively. Fractionated cells were pelleted, resuspended in medium, and spotted onto slides for *in situ* hybridisation as above.

### *Single Cell Clonal Assay and Statistical Analysis*

Single immature erythroid cells from Percoll gradient fractions were obtained by limiting dilution and cultured in 50 µl conditioned medium with erythropoietin and stem cell factor, for 3 days. For the methylcellulose assay, 5,000–10,000 cells were mixed with 1ml of methylcellulose solution in cfu- or bfu-conditioned medium and grown for 3 days at 37°C in 10% $CO_2$. Single colonies were fixed for primary transcript *in situ* hybridisation on heavy teflon supercured coated slides (Nutacon), ten wells per slide, each well treated separately to prevent mixing. α- and β-globin allelic expression patterns in all cells in each well were counted. The predicted distribution was derived from the following formula: (n!/(n!-x!*x!))*(0.8n-x*0.2x), where n = 30 (total number of clones) and x is the number of clones

predicted with a specific allelic expression pattern. The probability for each allelic distribution observed per clone was calculated using the npar $\chi^2$ test of the SPSS9 package. The distribution of probabilities observed was statistically tested against the expected distribution using the one-sample Kolmogorov-Smirnov test. The observed distribution was statistically different at p level p = 0.00. In graphing the data, we grouped the percentage of clones with all cells expressing 2α2β genes (100%), then the percentage of clones with 90% of the cells expressing 2α2β with the other 10% of cells having another transcription pattern (almost always just one other pattern), followed by the percentage of clones with 80% 2α2β cells, followed by the percentage of clones with 70% 2α2β cells, and finally all the clones that had 60% or less 2α2β-expressing cells.




| | Expressing | 1α | 2α | 2α1β | 2α2β | 1α1β | 1α2β | 2β | 1β |
|---|---|---|---|---|---|---|---|---|---|
| | Alleles α<br>Alleles β | | | | | | | | |
| C | % cells with primary transcripts<br>sdv | 4.03<br>0.97 | 5.31<br>1.84 | 8.69<br>0.53 | 73.19<br>2.63 | 3.66<br>0.54 | 4.94<br>1.08 | 0.09<br>0.19 | 0.09<br>0.09 |
| D | % cells with primary transcripts (+sdv) after fractionation | 0.21<br>0.30 | 0.00<br>0.00 | 3.80<br>1.16 | 84.57<br>0.76 | 3.83<br>0.03 | 6.77<br>0.66 | 0.84<br>0.59 | 0.00<br>0.00 |
| E | % cells with different α/β mRNA ratios | Rα/β > ctrl = 15.7 (Rα/β > 2 = 5.2) | | | Rα/β = ctrl 73.3 | | Rα/β < ctrl = 11 (Rα/β < 2 = 4.2) | | |

***Figure 1:*** *Primary Transcript In Situ Hybridisations of 14.5dpc Foetal Liver Cells*—(A and B) Two representative hybridizations with DNP-labeled mouse α-globin intron probe (red) and DIG-labeled mouse β-globin intron probe (green). (A) One focal plane. (B) Stack of confocal images. Different allelic expression combinations are visible. Hybridization efficiency was measured using probes against the introns and the 3rd exon of the α- or β-globin genes. Less than 1% of the exon-positive cells failed to show preceding exon or intron signal, indicating ~99% hybridization efficiency (data not shown). (C–E) Summary of α- and β-globin expression patterns in the nucleus and cytoplasm of 14.5dpc foetal liver cells. (C) α- and β-globin primary transcript patterns counted and tabulated as a percentage of total cells (n = 1200). (D) Re-counting of α- and β-globin primary transcript signals in fractionated mature erythroid cell fraction (n = 500). (E) Ratios of α/β cytoplasmic mRNA levels per cell (n = 200) were plotted as histograms (Figures 2C–2E) after normalization of the control (α/α or β/β) to a ratio of 1.0. Abbreviations: sdv, standard deviation; Rα/β, α/β cytoplasmic mRNA ratio; ctrl, control (α/α or β/β).

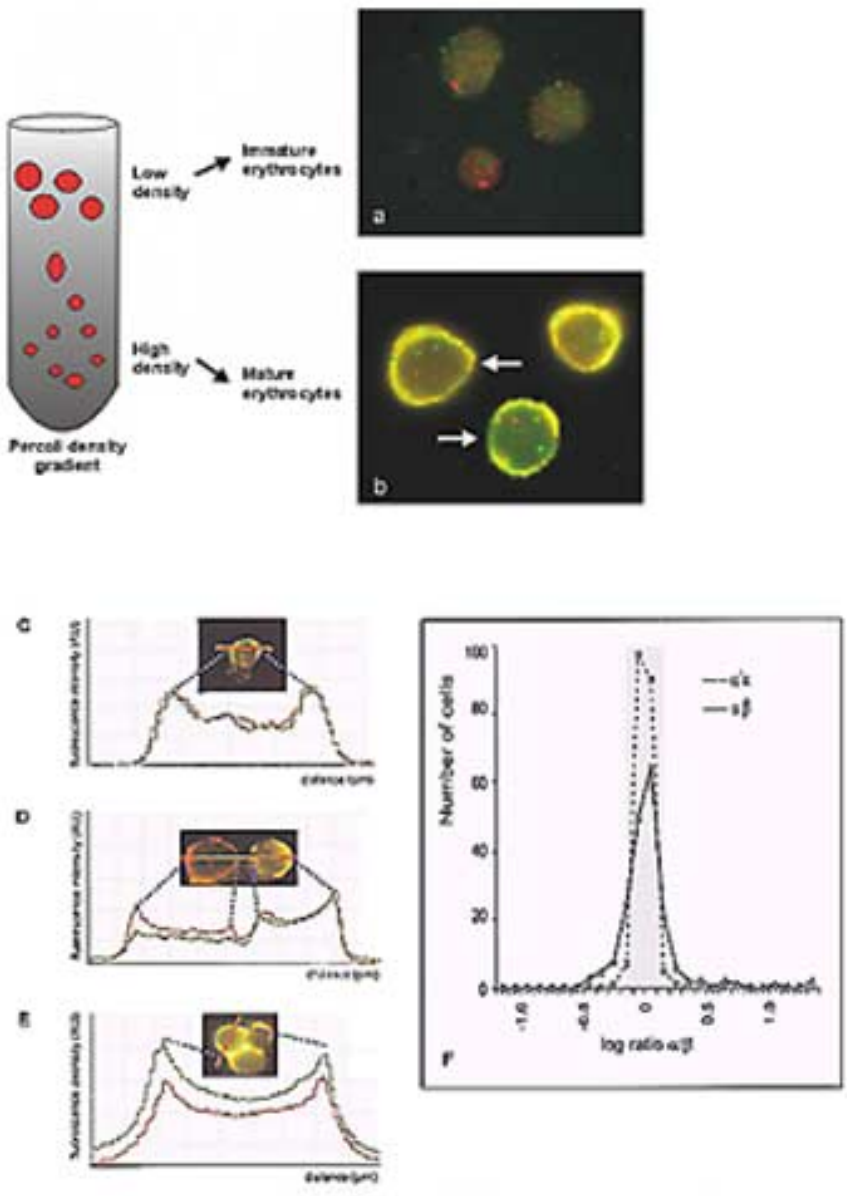



***Figure 2:*** *Visualization of Nuclear and Cytoplasmic RNA in Fractionated Erythroid Cells* —(A and B) Percoll density gradient size fractionation of erythroid cells. (A) Fraction enriched in large immature erythrocytes. Only α-globin nuclear staining is observed with no or very little cytoplasmic staining. No β-globin signal is detected. (B) Fraction enriched in small mature erythrocytes. Both α- and β-globin nuclear transcripts and cytoplasmic staining are observed (arrows). (C) Control mRNA *in situ* hybridisation using a mixture of one α-globin exon probe, half of which was labeled with DNP (red) and the other half with DIG (green). The two signals were overlaid and cytoplasmic intensities measured by drawing a line through single cells (n = 200) and plotted as two peaks in a histogram, one for the green signal (green line) and one for the red signal (red line). The average of two measurements was calculated. The ratio obtained served as control for background fluorescence and for the distribution intrinsic to the 1:1 ratio of red/green signal for the same mRNA. (D and E) Quantitation of signal intensities for α- (red) and β-globin (green) mRNAs. Three examples of cells with different α/β cytoplasmic ratios are shown. (F) Line representation of the ratios of α/β-globin mRNA in the cytoplasm as obtained from (C–E). The dotted line indicates the control distribution of α/α ratios, with the shaded area indicating cells with a 1:1 cytoplasmic mRNA ratios (β/β ratios are identical). 200 cells were scanned in randomly chosen fields. The ratios for signal intensities (α/α and α/β) were normalized to 1 using all ratios between 0.5 and 2.0. The cells were grouped with a ratio interval of a log of 0.1, and the number of cells falling in each group was plotted against the log of the ratios.

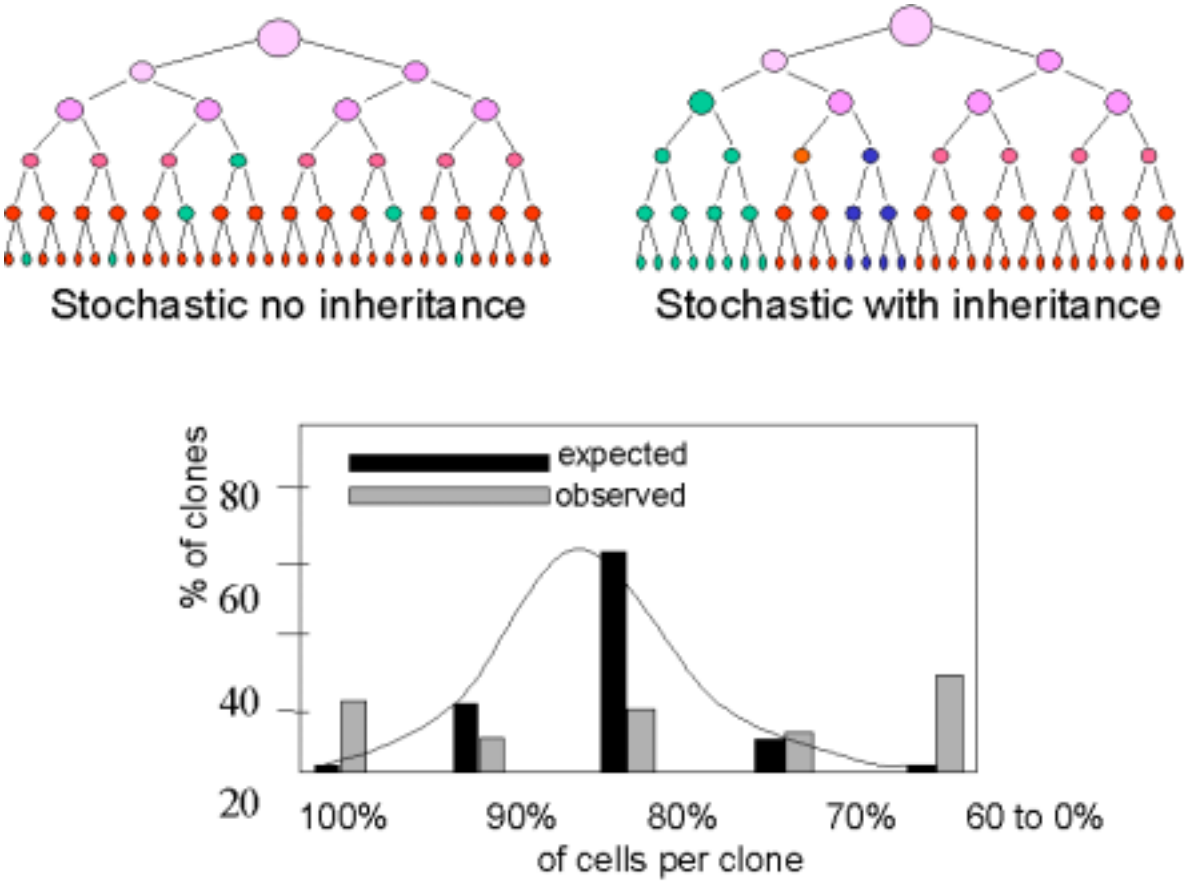



***Figure 3:*** *Clonal Analysis of Mouse α- and β-Globin Primary Expression Patterns*—(Upper panels) Examples of expected primary expression patterns if stochastic decisions are taken at the transcriptional activation step (left panel) or at a step prior to that with fixation of a choice in daughter cells (right panel). White, uncommitted cells; black, cells expressing $2\alpha2\beta$; gray, cells expressing a pattern other than $2\alpha2\beta$. (Bottom panel) All cells in each clone were scored for expression, and the percentage of clones was plotted against the expression pattern per clone. These patterns were grouped as described in Experimental Procedures. Black bars and curve represent expected distribution for transcription patterns made according to (left upper panel). Gray bars are the observed distribution, suggesting fixing (Upper right panel) around the 3rd to 4th replication of cells.

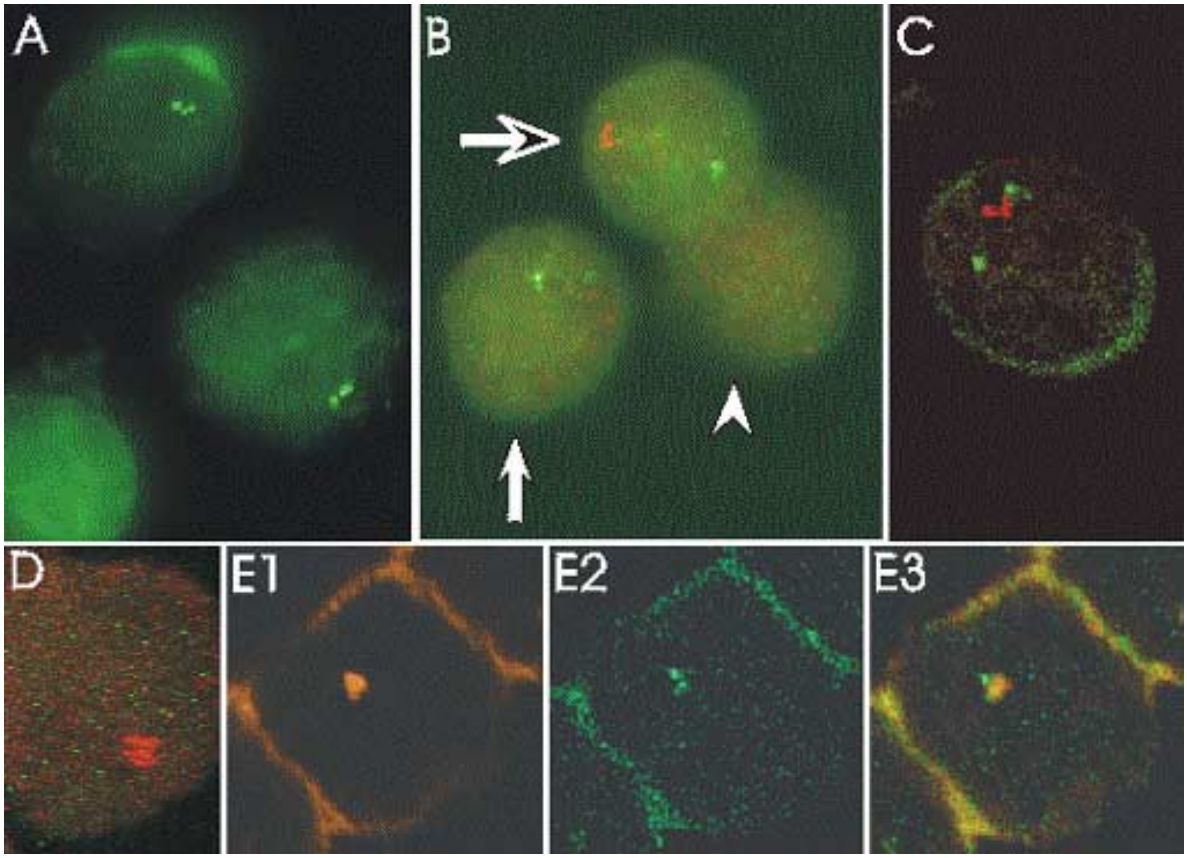



***Figure 4:*** *In Situ Hybridisation of Foetal Liver Cells from a Transgenic Mouse Heterozygous for a Three-Copy Human β-Globin Locus*—(A) Hybridization with a 12 kb LCR probe (green). (B and C) Hybridization with probes for α (green) and β (red) introns detecting nuclear primary transcripts. Photographs by CCD camera (B) or by confocal microscopy with two focal planes superimposed (C). (D) Deconvolution of confocal images as in (C) and rotation to visualize RNA tracks in one focal plane. (E1–E3) *In situ* hybridisation to detect human β-globin primary transcripts with an intron probe (green, [E2]) and the β-globin loci with a PAC probe (red, [E1]) to show colocalization (E3) of human β transcription and the human β loci by superimposing (E1) and (E2).