3. Results

Differentiation of the WEHI-231 line

AID is generally considered a marker of activated B cells in the germinal center (Muramatsu et al., 2000). Indeed, according to the NCBI database, AID mRNA expression is limited to cells of the germinal centers. WEHI-231, a commonly used cell line is classified as an immature B cell line. However, we found expression of AID in the WEHI-231 progeny line FS and its subclone HM (see below). In contrast to FS and HM, the original WEHI-231 line does not express the canonical AID. Instead it expresses an 18-kD protein that may or may not be an AID variant (see below). In addition, we found expression of IgD—a mature B cell marker—in the original WEHI-231 line CP, whereas FS and HM do not express IgD (see below). Based on these results we can describe the evolution of the WEHI-231 cell line as follows: The original WEHI-231 line is a mature B cell line. Upon activation, this cell line lost expression of IgD and the 18-kD band, but retains expression of IgM and now synthesizes the canonical AID at 24 kD. This activated cell is a precursor for the HM clone, which we selected for low surface IgM expression, but it also gave rise to the FS line. Presumably due to the expression of AID, FS switched the C-region of IgM of the a allotype, which is expressed in all other WEHI-231 lines and HM, to the b allotype, and HM lost surface IgM upon accumulation of mutations in the κ chain. The differentiation stages of B cells and WEHI-231 are summarized in Fig. 6.
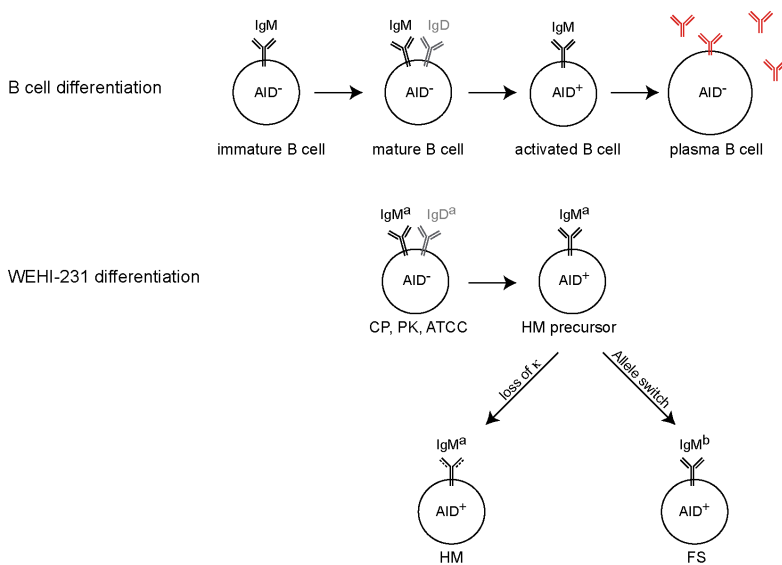

Fig. 6. Comparison of differentiation stages and ontology of primary B cells and WEHI-231. Upper row, B cell development. Lower row, differentiation of the WEHI-231 line. In red, secreted antibodies of an Ig class other than IgM.

FS expresses AID

Fig. 7A compares the expression of cDNA encoding AID in the activated B cell line 18-81 (lane 3), a cell line that hypermutates and switches its Ig genes (Burrows et al., 1981; Li et al., 2003; Meyer et al., 1986), FS (lane 4), its subclone HM (lane 5), and the pre-B cell line 70Z (lane 6), which is active in neither SHM nor CSR. Surprisingly, while the 70Z does not express it, the FS cells express AID at levels comparable to that of 18-81.
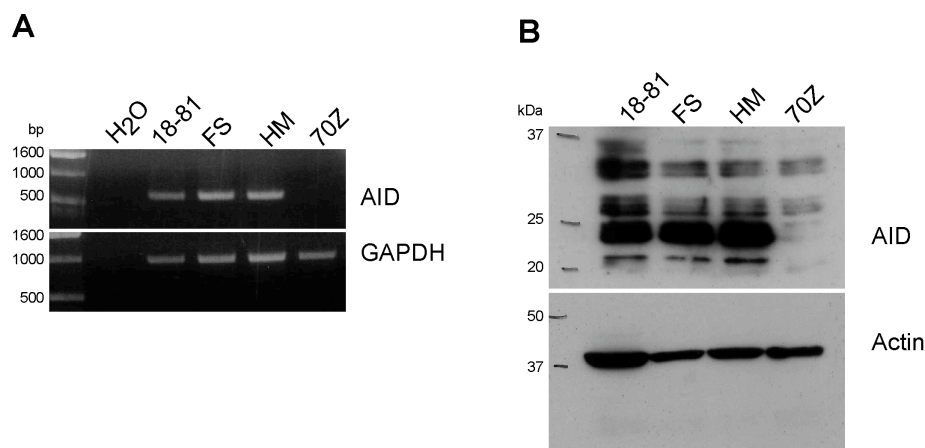

Fig. 7. AID expression in the cell line FS and its subclone HM. (A) RT-PCR amplification of AID transcripts, and GAPDH transcripts as mRNA amplification and loading controls. H₂O, amplification without template, but with all the other reagents. Size of amplified AID segment: 596 bp. (B) Western blot assaying for AID expression; actin as loading control.

Because WEHI-231 has been classified as an immature B cell line, we thought that RNA expression at the "wrong" differentiation stage might be accompanied by mutations in the gene encoding AID. However, the fragment sizes of the amplified cDNAs did not show any, nor did the 900 nt cDNA sequence from the FS line, or from its subclone HM, contain any mutation (not shown). Also, the Western blot with antibodies against AID detected the predicted 24-kD band in FS, as in the 18-81 line (Fig. 7B, lanes 1-3). No such band is seen in the 70Z line (Fig. 7B, lane 4). The AID protein levels are comparable in the FS clones and 18-81; they may be even higher in FS.

AID variant in WEHI-231 precursor lines

Because we wanted to know whether or not AID is generally expressed in WEHI-231, we compared FS and HM to 'early' freezings of the original WEHI-231 line, in the following denoted as precursors. We found that the precursor WEHI lines PK, CP and ATCC do not express the canonical AID band at 24 kD. Instead there is an 18-

kD band (Fig. 8). Because of the precursor-progeny relationship between lines ATCC and FS, an 18-kD band most likely does not represent an aberrant AID protein, but an alternative (inactive form). Apparently, the WEHI-231 seems to have evolved in culture to express AID, which opens up the possibility to study both AID induction and putative AID variants.
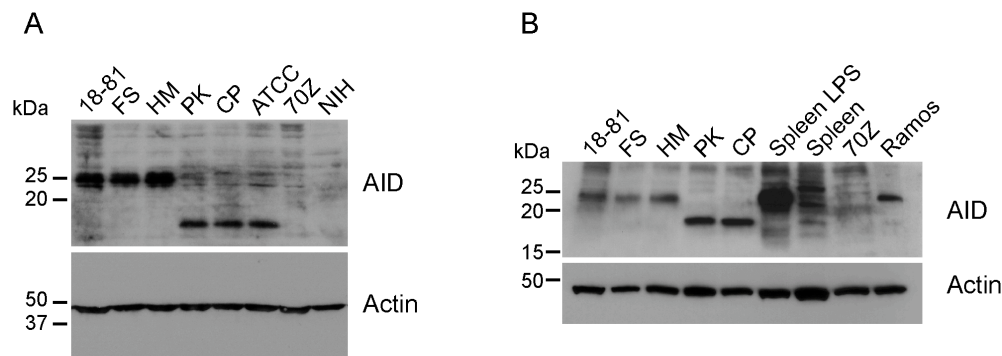

Fig. 8. AID expression in WEHI-231 lines, control lines and (activated) spleen cells. Cell lysates were electrophoresed in SDS gels, and the gels were Western blotted and developed with antibody to AID (above) or actin (below). Numbers to the left of the blots indicate molecular mass (in kD). 70Z, pre-B cell line 70Z/3; NIH, fibroblast line; Ramos, human B cell line; spleen, spleen cells; spleen LPS, spleen cells stimulated with LPS.

Quantification of AID in WEHI-231 lines

Although we could detect the canonical AID protein in the precursor lines PK, CP and ATCC, it seemed to be expressed at very low level (Fig. 8). We therefore wanted to quantify expression levels and performed Taqman analysis with primers and probes priming in exons 4-5 of AID. Indeed, there is very little AID transcript expressed in the precursor lines. Unexpectedly, the 18-81 line and the LPS activated spleen cells have the same steady-state levels of AID cDNA, whereas there is even 3.5 more AID transcripts in the FS line than in the activated spleen cells (Fig. 9). The amount of AID transcript (Fig. 9) does not reflect the amount of AID protein expressed (Fig. 7, 8)—with by far the most AID protein expressed in the spleen cells—suggesting that there is translational or post-translational regulation involved.
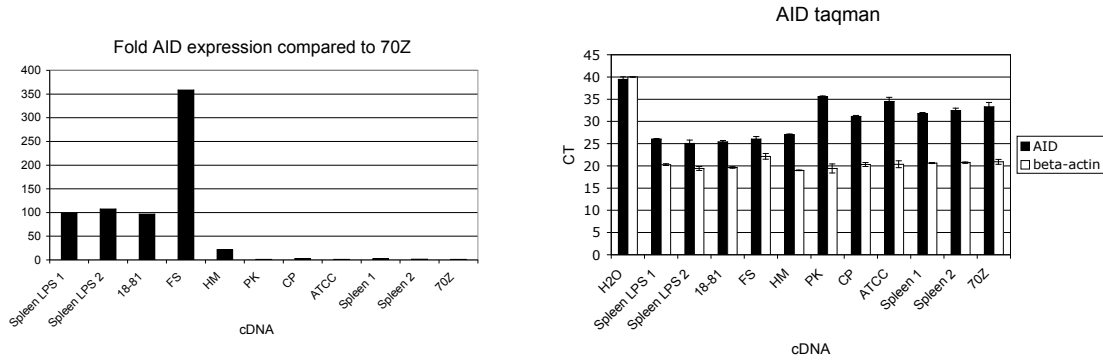
Fig. 9: Taqman of AID cDNA from WEHI-231 lines. Primers are located at the border of exon 4 and 5. The probe is spanning the junction between exon 4 and 5. Spleen, spleen cells. Left: data normalized to beta-actin and H_2O (without cDNA) control. Right: non-normalized data. Spleen LPS 1 + 2, two independent cultures with 3 x 10^5 cells/ml LPS stimulated; FS and HM, WEHI-231 progeny lines; PK, CP, and ATCC, precursor lines. 70Z, non-AID expressing. N=3.

AID KO mouse

In most experiments on canonical AID expression we used the pre-B cell line 70Z as a negative control. To prove that the 18-kD band is an AID form, we also included AID KO cells, as a stricter negative control. We would identify the 18-kD band as AID, if it was absent in lysates from AID KO spleen cells. Because less than 50 % of the leucocytes of a spleen are B220 (a B cell marker) positive (Gilman et al., 2002), spleen cells were enriched for B220 positive cells by FACS, followed by protein extraction. But we did find the 18-kD band expressed in non-stimulated AID KO spleen cells, as well as in non-stimulated WT cells (Fig. 10). Upon stimulation with LPS, the WT expresses the canonical 24-kD band, whereas the AID KO shows only a very faint band at the same height. This result suggests that the 18-kD protein may not represent an AID form, but another protein that is specifically expressed in non-activated B cells.
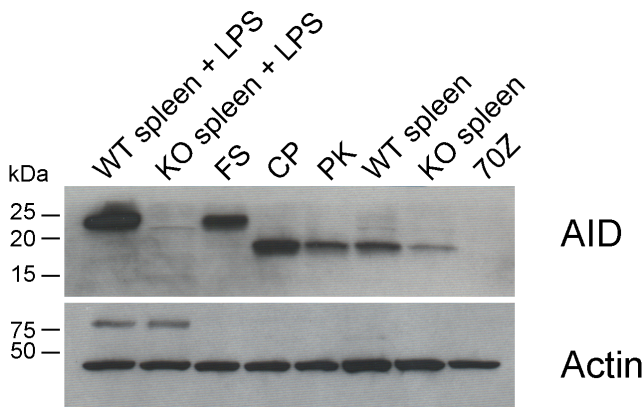

Fig. 10. AID Western blot with lysates from BALB/c wildtype (WT) and knock-out (KO) spleen cells. Cells were stimulated (spleen + LPS) with LPS or untreated (spleen). FS, canonical AID positive control; CP, 18-kD band positive control, 70Z canonical AID negative control.

We then wanted to exclude the possibility, that the 18-kD band is due to leakage of the KO construct, i.e., mRNA encoding part of the AID gene would still be generated. We therefore designed oligonucleotides priming in the Neo cassette, which disrupts exon 2 and part of exon 3 of AID (Fig. 11A), and exon 5 of AID. Indeed, we found transcripts with this primer combination in the LPS-treated and non-treated KO cells. To verify the identity of these transcripts as AID encoding, we gel extracted and sequenced these; and found that they represent various splice forms of Neo-AID transcripts (Fig. 11B), schematically represented in Fig. 11C. Nevertheless the 18-kD band is probably not encoded by these transcripts. The Neo-AID transcription is controlled by the Neo promoter and thus constitutively expressed, and is transcribed regardless, whether in presence or absence of LPS (Fig. 11B). Therefore, the 18-kD band should be present in both, LPS-treated and non-treated spleen cells, which is not the case. However, the presence of these (in part) correctly spliced transcripts makes it more likely that there are also transcripts off the endogenous AID promoter(s), which may splice out the neo cassette and encode polypeptides other than the canonical 24-kD AID protein.
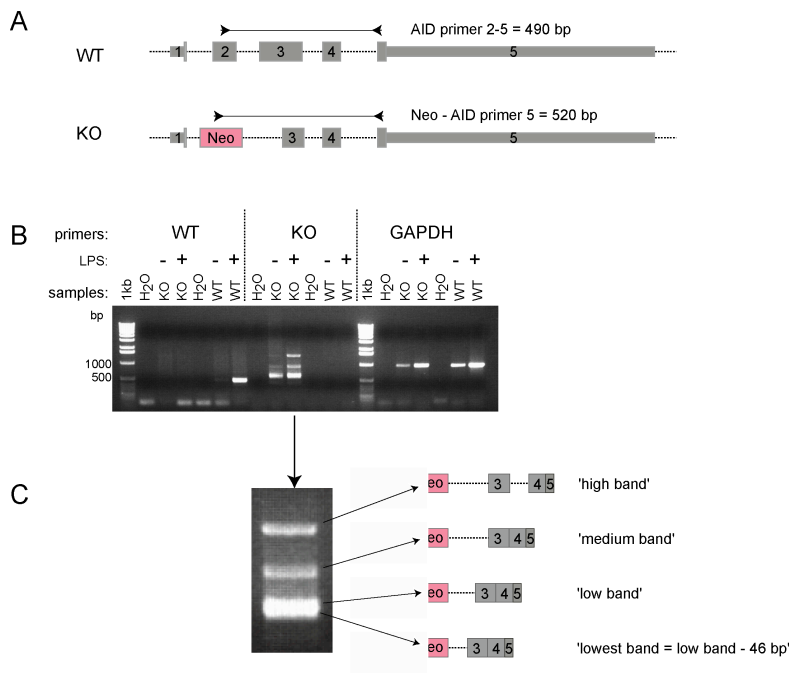
Fig. 11. Neo-AID transcripts in AID KO mice. (A) Schematic representation of AID encoding alleles, and primer design for WT and KO mice. (B) Amplification products with primers shown in (A), either with WT 5' primer, or with KO 5' primer, or with GAPDH primers. Treatment with (+) or without (-) LPS; lower row, cDNA template used for amplification. (C) Magni-fication of amplification products from LPS treated KO spleen, and schematic representation of the various Neo-AID splice variants.

AID encoding transcripts

Because of the various AID exon-containing transcripts in the AID KO mouse, there remains the possibility that the 18-kD band is an AID precursor. A transcript containing a premature stop codon may encode such a precursor, or may form secondary structures at the 3' end that result in inhibition of translation. Upon activation of the B cell the AID precursor may be edited and the canonical AID will be expressed. This AID precursor transcript must not contain the neo cassette, because the cassette contains a premature stop codon. In addition, it has to include the translated part of AID exon 5—the part of AID that encodes the epitope of the antibody to AID.
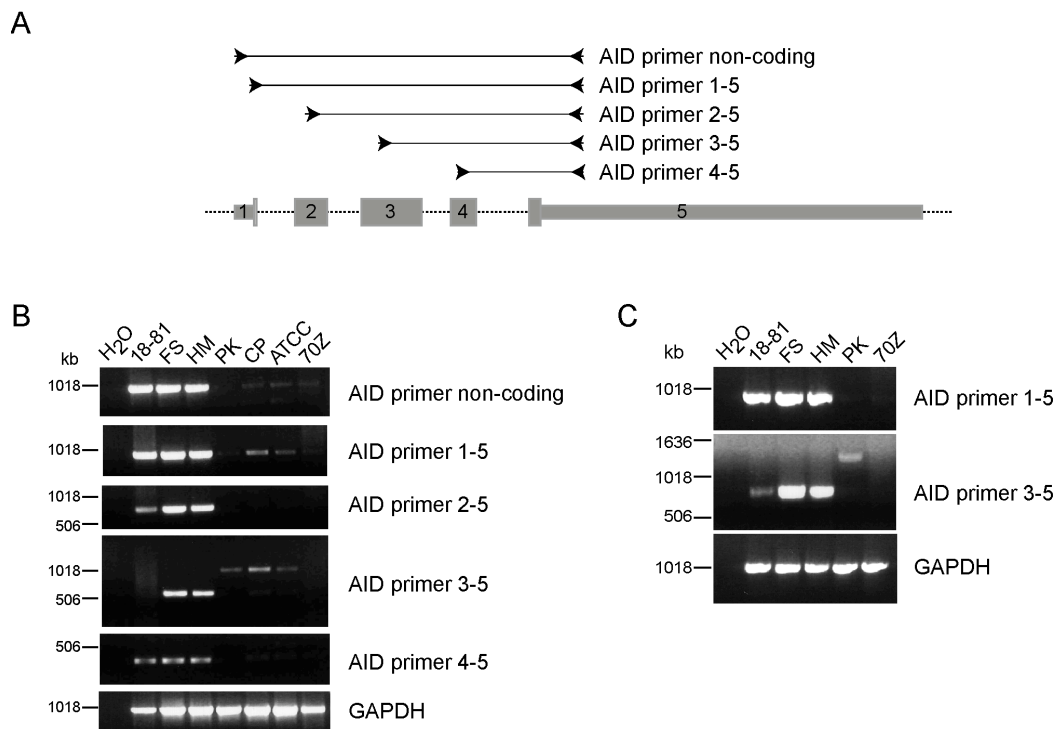
Fig. 12: AID cDNA expression in lines WEHI-231 lines and 18-81. (A) Schematic representation of the primer pairs used in amplification. Shown are the five exons of AID; smaller boxes: untranslated part of the canonical protein. (B) Agarose electrophoretic gels of amplification products obtained with diverse primer pairs for AID and GAPDH. H_2O, no cDNA; 18-81, mutator and switch-positive cell line with endogenous AID expression; FS and HM, late WEHI-231 and its clone, respectively; PK, CP, ATCC early freezing of the WEHI-231 line; 70Z, 70Z/3, an AID negative line. (C) Repeat with some of the samples and primers.

In order to find the transcript encoding for the 18-kD band, cDNA with oligonucleotides priming in various AID exons was amplified (Figs. 12A). In the precursor WEHI-231 lines we found little AID transcript with all different primer combinations. With the primer 'non-coding' and '3-5', there were some additional

bands that were specific to the precursor lines, but sequencing of these bands showed that they were the result of non-specific priming of the oligonucleotides. Thus, at the moment we do not have in hand a cDNA encoding the 18-kD band.

However, there was a difference in the RT-PCRs between 18-81 and FS/HM: in 18-18 there is only a small amount of amplification product with the primer pair 3-5, although the primer pair 1-5 amplifies the canonical fragment perfectly well (Fig. 12B and C). This is quite intriguing, as the 18-81 line shows hypermutation of exogenous substrates and endogenous H locus at a rate that has been estimated to be near to the one of the hypermutation process in *vivo*, and it switches its H chain efficiently from μ to γ2b (only). Yet as mentioned above, both 18-81 and progeny WEHI-231 lines seem to express the canonical 24-kD AID protein.

We bacterial-subcloned and sequenced the amplification products of the primer combinations '1-5' and '3-5' from 18-81 (10 sequences), CP (4 sequences) and PK (3 sequences). Besides the canonical AID, we found two splice variants: a 30 bp deletion of exon 4 was found in 18-81 and CP; and a deletion of the whole exon 4 in 18-81 (Fig. 13). These splice variants had been already found in humans (Albesiano et al., 2003; McCarthy et al., 2003; Noguchi et al., 2001). We can be sure that our sequences are derived from the respective 18-81 and WEHI-231 lines. This is because we found an allelic difference in exon 5 at position 223 of the untranslated region, where the a allele has a G and the b allele an A. The reference sequence for the canonical AID sequence (accession number: AF132979) is derived from a C57BL/10 mouse, and therefore likely of the b allotype. Indeed, in the WEHI-231 lines, because there are derived from a BALB/c x NZB F_1, had half of the clones of the b allele and the other half of the a allele. As expected, all sequences from 18-81—a line derived from a BALB/c mouse—were of the a allotype.

None of the cDNAs cloned can explain the pronounced decrease in amplification with the primer pair 3-5 in 18-81. Nor did the subcloning give us any further hint as of the identity of the 18-kD band. At the time of the writing of this thesis, various cDNAs encoding for AID are being cloned and sequenced from precursor and progeny lines and from small and activated B lymphocytes.
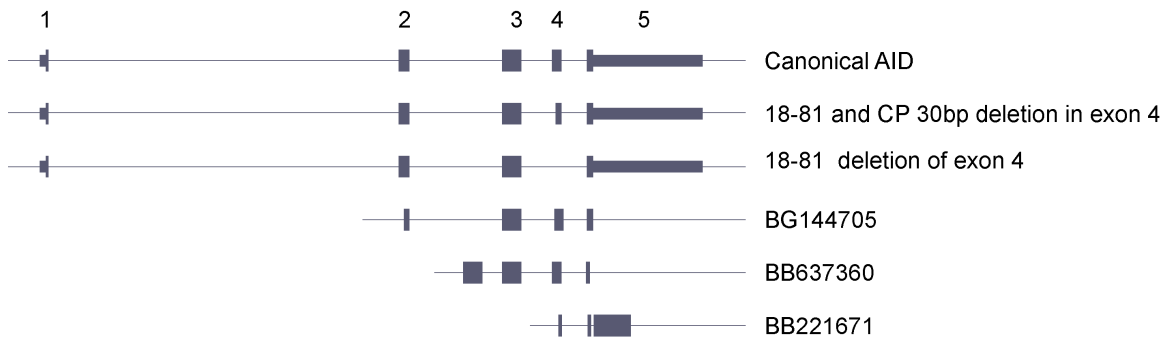
Fig. 13: Splice variants and ESTs encoding AID from the UCSC database and from the 18-81 cell line. Numbers above boxes, exons 1 to 5; smaller boxes, untranslated regions.

Because the rabbit antibody used for Western blotting (Fig. 7, 8) was raised against a peptide of the translated part in exon 5, we use this sequence as an anchor to obtain the transcript encoding for the 18-kD band (work in progress). These experiments are designed to test whether the 18-kD band represents translation from an alternative splice form. But one should also keep the possibility in mind, that the 18-kD band may be a degradation product, or a polypeptide that cross-reacts with the antibody to AID.

IgD expression in WEHI-231

Because AID is a marker for activated B cells, it was surprising to find it expressed in the WEHI-231, although this line is known as an immature B cell line. However, we only found canonical AID expression in the WEHI-231 progeny lines FS and HM.
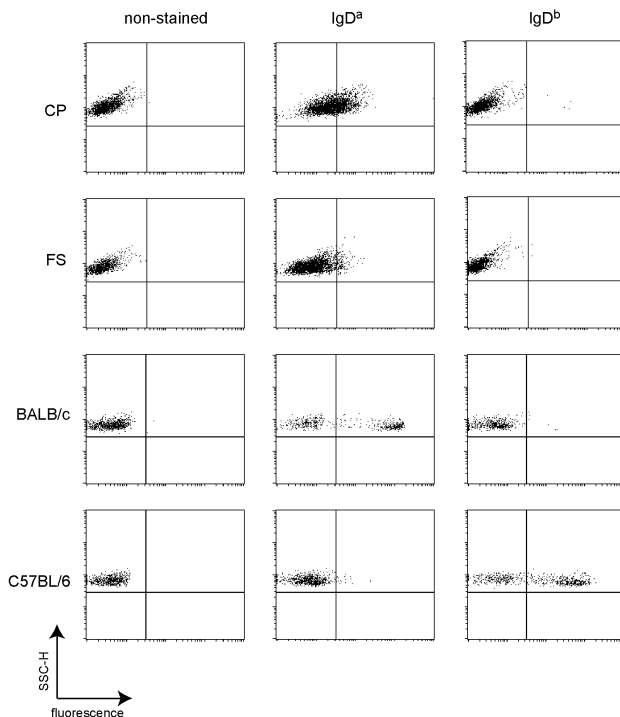
Fig. 14. Expression of IgD of a and b allotype in WEHI-231. CP, WEHI-231 precursor line; FS, WEHI-231 progeny line; BALB/c, spleen cells from a BALB/c mouse as a positive control for the a allotype, C57BL/6, spleen cells from a C57BL/6 mouse as a positive control for the b allotype; Left colum, non-stained; middle column, cells stained with anti-IgD[a]-FITC; right column, cells stained with anti IgD[b]-PE antibody, X-axis, fluorescence intensity on a logarithmic scale; Y-axis, side scatter.

In contrast, the WEHI-231 precursor lines CP, PK and ATCC do not express the canonical 24-kD AID, but instead an unidentified 18-kD protein, which is specific for non-activated, mature B cells. This would classify the precursor lines as mature B cells, and the progeny as activated B cell. To confirm, we stained for IgD—a marker for mature B cells. Indeed, we found that CP is positive and FS is negative for IgD (Fig. 14). Based on the expression of IgD, canonical AID and the 18-kD band we re-classify WEHI-231 as a mature B cell line and it progeny lines FS and HM as activated B cell line.

The FS line expresses nuclear UNG

Because enzymatic activity of AID may be followed by uracil-DNA glycosylase (UNG) activity in class switching and hypermutation (Imai et al., 2003; Li et al., 2003; Rada et al., 2002b), we wanted to assess the integrity and the expression level of the gene encoding this enzyme. There are two splice variants, UNG1 and UNG2, which localize to the mitochondria and to the nucleus, respectively (Nilsen et al., 1997; Otterlei et al., 1998). The nuclear form, UNG2, takes part in CSR and SHM. Using RT-PCR, we determined that FS expresses mRNA of normal size and no lower steady state levels than the cell line 18-81, 70Z/3, or activated splenic B cells (Fig. 15A). Because the FS line was derived from a cross between the two inbred strains BALB/c and NZB, one might expect allelic forms of the enzyme. Indeed, when we sequenced the cDNA encoding UNG2, we found two different alleles. The a allele is as present in the cell line 70Z/3 and is 100% identical to the sequence Y08975 in the GenBank; the other is as in the 129Sv mouse strain and, except for a T at position 796 (a silent mutation), which is shared with Y08975, is identical to the sequence AF174485 (GenBank). Apart from the differences in the non-coding regions (non-shaded in Fig 15B), the two alleles differ by one mutation in the coding region (shaded in Fig. 15B). At position 829, the a allele has a T, where the b allele has a C; this changes the amino acid residue from tyrosine to histidine (Fig. 15B). We do not know whether this changes the activity of the enzyme; but because they are present in standard mouse strains, we assume that at least one of these alleles is functional, and so we further assume that there is functional UNG2 enzyme present in the FS line.
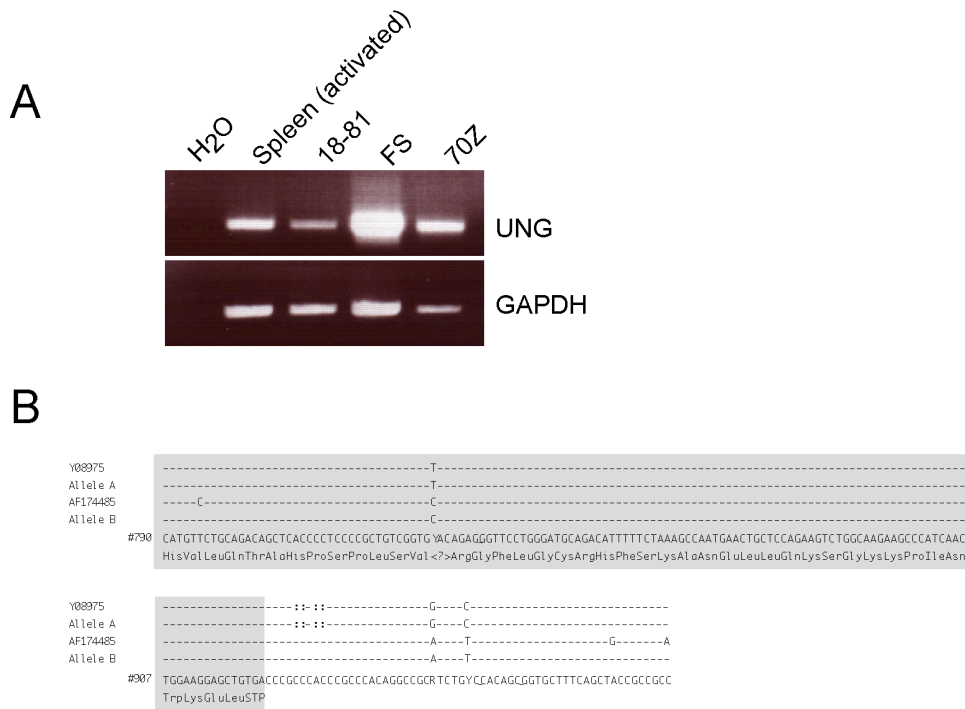
Fig. 15. (A) RT-PCR amplification of UNG2 transcripts; and GAPDH transcripts as loading controls. Size of amplified UNG segment: 997 bp. (B) Sequences of UNG2 alleles in FS. Although the complete cDNA encoding UNG was sequenced, only the part of the sequence starting with nucleotide #790 is shown here.

FS expresses Ku70/Ku80

The Ku70 and Ku80 proteins of the non-homologous end joining repair pathway are both required for switch recombination (Casellas et al., 1998; Manis et al., 1998). Using RT-PCR, we determined that FS expresses mRNA of apparently normal size and no lower steady state levels than the cell line 18-81, or 70Z/3 (Fig. 16). Furthermore, the DNA sequences of the Ku70 and Ku80 alleles did not contain any mutations. It is therefore likely that Ku70/Ku80 function normally in this cell line.
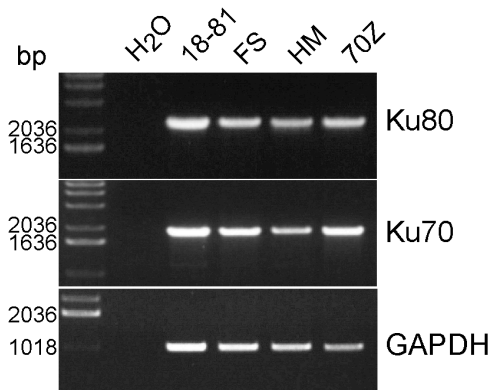
Fig. 16. Ku70/Ku80 transcripts; sizes of amplified segments: Ku70, 1905 bp; Ku80, 2248 bp; GAPDH, 983 bp.

Canonical mutations at the H locus

As we found no mutations in AID or UNG2 that would render these enzymes non-functional, we assayed the V region gene segments at the H chain locus for hypermutation. Because the H alleles (and the L alleles) of WEHI-231 were not known, we amplified the V region and Cμ1 of FS and its subclone HM with a generic primer for the V-region via RT-PCR, subcloned the amplification products and sequenced them. We then designed a primer specifically for the V region. This primer primes 361 nucleotides 5' of the V exon. Fig. 17A and 17B shows the sequences of the two alleles, from genomic DNA for the active allele, and from cDNA for the silent allele. Underneath the nucleotide sequences are given the amino acid sequences. The shaded area represents the D region, after which the J_H region needs to be in frame for the allele to contribute to surface Ig. This is the case for the sequence in Fig. 17A, which therefore is the active allele; the J_H region in the sequence of Fig. 17B is out of frame, which therefore represents the silent allele. Although not perfect, the best fit for the active allele corresponds to nucleotide position 123800 to 124080 in the Ensembl database; for the silent allele to position 46223 to 46501.

Fig. 17. Sequences of the VDJ and Cμ1 gene segments of active (A) and silent (B) alleles. In (A) and (B), the most similar germline sequence is given above the translation into polypeptide; dashes indicate identity of the FS sequence with the sequence in the database. The shaded box indicates the D region. DNA and cDNA sequences were determined for both alleles, A and B, but only A DNA and B cDNA are shown.

To trace endogenous mutations, we compared the sequences of a subclone HM to the ones of the cell line. Subcloning the amplified DNA into bacteria preceded the sequencing of the V region from genomic DNA. In this way, the individual V region in different cells, and not a composite of all possible sequences, was obtained. We found three mutations between the five sequences of FS and five sequences of the subclone HM. The T to G mutation at position 235 occurred in one of the cells of the line. This mutation introduced by the method. However, the other two mutations at positions 290 and 406 are not method induced, because they are present in all five sequences, and they were present when cDNA was amplified with Taq polymerase instead of Pfu. At position 290, in the D region, we found a G to A (the germline sequence has a G) at position 406, an A to G (the germline sequence has an A).
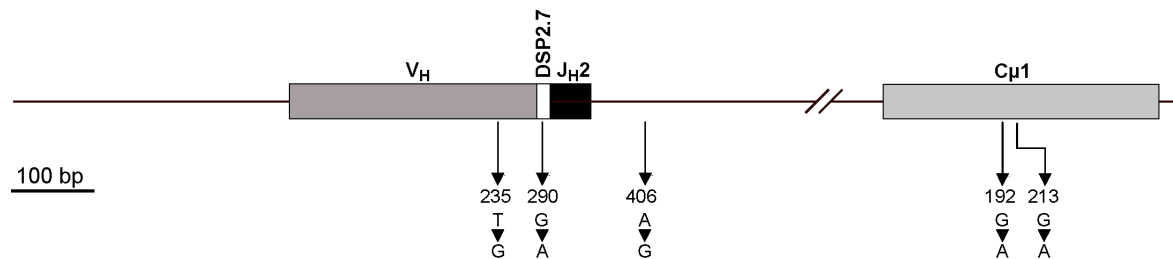
Fig. 18. Summary of the somatic mutations between cell line FS and subclone HM.

Because it is obvious that in general mutations are much less abundant in the C region, we were quite surprised to find two mutations in the Cμ1 region (Fig. 18), while surveying 12 sequences, 6 from FS and 6 from HM. Both nucleotide substitutions are silent mutations, and therefore probably not selected for. One mutation, from G to A, embedded in a RGYW motif, occurred at position 192 in one of the six cells of the uncloned cell line (Fig. 18). As it occurred only once, it might have been introduced by the amplification. But the other mutation, from G to A, outside of an RGYW motif, was found at position 213 in all sequences of the uncloned cell line and thus for sure has occurred in the cell line. The Cμ a allele is distinguished from the b allele by monoclonal antibodies through a single codon change, at nucleotide position 292 of exon Cμ1, where the b allele has an A, and the a allele has a G (Schreier et al., 1986). Arginine, encoded by AGA, in this position prevents binding of the anti-μ[b] antibody MB86 (Nishikawa et al., 1986). In the six sequences of the subclone, three have the lysine encoding AAA codon (the b allele), and three have the AGA codon (the a allele). Out of the 6 sequences of the uncloned line, none represented the a allele. At position 213, all six subclone cells (i.e., a and b alleles) have a G, as does the germline sequence. Therefore, it seems likely that the

cells of the uncloned line surveyed underwent a mutation, from G to A. This would imply that a majority of the cells of the FS line are derived from a single mutant that outgrew most of the sister cells, some of which, including the founder cell of subclone HM, have retained the germline nucleotide. At position 192 one of the six FS sequences bears a G to A mutation; this mutation could be method introduced.

The κ locus

We determined the κ V region in FS, encoded by the Vκ cs1 and Jκ1, using RT-PCR with a 5' generic Vκ and a 3' Jκ5 primer. We sequenced the κ promoter region of FS and HM, but did not find any mutations. We then amplified κ from all WEHI-231 strains with primers in the leader and the 3'UTR. Fig. 19 shows the amplification products. All WEHI-231 seem to express κ mRNA at the same level, which will be important for the experiments described later.
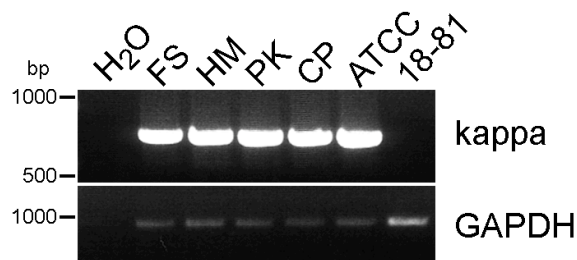
Fig. 19. Amplification of κ chain of all WEHI-231 strains; 18-81 as a negative control.

We subcloned and sequenced the amplification products shown in Fig. 19. We have sequenced 5-6 subclones for each strain and found many mutations when we compared the κ chain sequences between the WEHI-231 strains. Four mutations are present only in HM and are labeled (*) in Fig. 20. Three of these mutations result in two amino acid exchanges: the first two mutations are located in the complementarity-determining region 1 (CDR1) and change the codon TAT (tyrosine) to TGG (tryptophan) and the third coding mutation is located in the framework region 2 (FR2) and changes CTC (leucine) to TTC (phenylalanine). The fourth, non-coding mutation is located in Jκ1 and mutates CTG to TTG. Because the mutations in FR2 and Jκ1 are C to T, with the C embedded in a WRCY motif, they may have been introduced by AID. Mutations not labeled were found in single clones only, and may or may not be method-induced. The Vκ reference sequence (NCBI accession number U16688) was derived from a BALB/c mouse (White et al., 1996). The reference sequence for the constant region (NCBI accession number BC092251) was derived from a FVB/N mouse strain.

```
                    LEADER                                                                                    CDR1
U16688  -----------------------------------------                              -----------------------------------------
   FS   -----------------------------------------                              -----------------------------------------
   HM   -----------------------------------------                              -----------------------------------------
   PK   -----------------------------------------                              -----------------------------------------
   CP   -----------------------------------------                              -----------------------------------------
 ATCC   -----------------------------------------                              -----------------------------------------
   #1   TGTTGTTCTGGATTCCTGCTTCCAGAGGTGATGTTGTGGTGACTCAAACTCCACTCTCCCTGCCTGTCAGCTTTGGAGATCAAGTTTCTATCTCTTGCAGGTCTAGTCAGAGTCTTGCAAACAG
         LeuPheTrpIleProAlaSerArgGlyAspValValValThrGlnThrProLeuSerLeuProValSerPheGlyAspGlnValSerIleSerCysArgSerSerGlnSerLeuAlaAsnSe

                CDR1                              FR2
U16688  -------------------------------------------A-------G-----A--G---------C-----------------------------------------
   FS   --**------------------------------------*-------T-----T-----------------A--G---------------------------------------
   HM   --GG------------------------------------------T-----------------------------------------------------------------
   PK   -----------------------------------------------------------------------------------------------------------------
   CP   -----G-----------------------------------------------------------------------------------------------------------
 ATCC   -----------------------------------------------------------------------------------------------------------------
  #125  TTATGGGAACACCTATTTGTCTTGGTACCTGCACAAGCCTGGCCAGTCTCCACAGCTCCTCATCTATGGGATTTCCAACAGATTTTCTGGGGTGCCAGACAGGTTCAGTGGCAGTGGTTCAGGG
         rTyrGlyAsnThrTyrLeuSerTrpTyrLeuHisLysProGlyGlnSerProGlnLeuLeuIleTyrGlyIleSerAsnArgPheSerGlyValProAspArgPheSerGlySerGlySerGly
          Trp                                                                                           Phe

                                                                              Jκ1
U16688  -------------------------------------------------------GTAC---------A--G--G-------------A---
   FS   -------------------------------------------------------------------------------------------
   HM   -------------------------------------------------------------------*------------------------
                                                                           T
   PK   -------------------------------------------------------------------------------------------
   CP   ----------------------------------------------------------------------G--------------------
 ATCC   ----------------------------------------G-----------------------------------------------
  #249  ACAGATTTCACACTCAAGATCAGCACAATAAAGCCTGAGGACTTGGGAATGTATTACTGCTTACAAGGTACACATCAGCCTCCGACGTTCGGTGGAGGCACCAAGCTGGAAATCAAACGGGCTG
         ThrAspPheThrLeuLysIleSerThrIleLysProGluAspLeuGlyMETTyrTyrCysLeuGlnGlyThrHisGlnProProThrPheGlyGlyGlyThrLysLeuGluIleLysArgAlaA

BC092251  >#1>                    ---------------------------------------------------------------------------------------------------------
     FS   -----------------------------------------------------------------------------------------------------------------------------
     HM   -------------------------------------------------------------------------------------------G-----
     PK   -----------------------------------------------------------------------------------------------------------------------------
     CP   -------G---------------------------------------------------------------------------------------------------------------------
   ATCC   -------G---------------------------------------------------------------------------------------------------------------------
   #373   ATGCTGCACCAACTGTATCCATCTTCCCACCATCCAGTGAGCAGTTAACATCTGGAGGTGCCTCAGTCGTGTGCTTCTTGAACAACTTCTACCCCAAAGACATCAATGTCAAGTGGAAGATTGA
          spAlaAlaProThrValSerIlePheProProSerGluGlnLeuThrSerGlyGlyAlaSerValValCysPheLeuAsnAsnPheTyrProLysAspIleAsnValLysTrpLysIleAs

BC092251  #94   ---------------------------------------------------------------------------------------------------------------------
     FS   ----------------------------G---------------------------------------------------------------------------------------------
     HM   -----------------------------------------------------------------------------------------------------------------------------
     PK   -----------------------------------------------------------------------------------------------------------------------------
     CP   -----------------------------------------------------------------------------------------------------------------------------
   ATCC   -----------------------------------------------------------------------------------------------------------------------------
   #497   TGGCAGTGAACGACAAAATGGCGTCCTGAACAGTTGGACTGATCAGGACAGCAAAGACAGCACCTACAGCATGAGCAGCACCCTCACGTTGACCAAGGACGAGTATGAACGACATAACAGCTAT
          pGlySerGluArgGlnAsnGlyValLeuAsnSerTrpThrAspGlnAspSerLysAspSerThrTyrSerMETSerSerThrLeuThrLeuThrLysAspGluTyrGluArgHisAsnSerTyr

BC092251  #218  ----------------------------------------------------------------
     FS   ----------------------------------------------------------------
     HM   ----------------------------------------------------------------
     PK   ----------------------------------C-----------------------------
     CP   ----------------------------------------------------------------
   ATCC   ----------------------------------------------------------------
   #621   ACCTGTGAGGCCACTCACAAGACATCAACTTCACCCATTGTCAAGAGCTTCAACAGGAATGAGTGTTAG
          ThrCysGluAlaThrHisLysThrSerThrSerProIleValLysSerPheAsnArgAsnGluCys STP
```
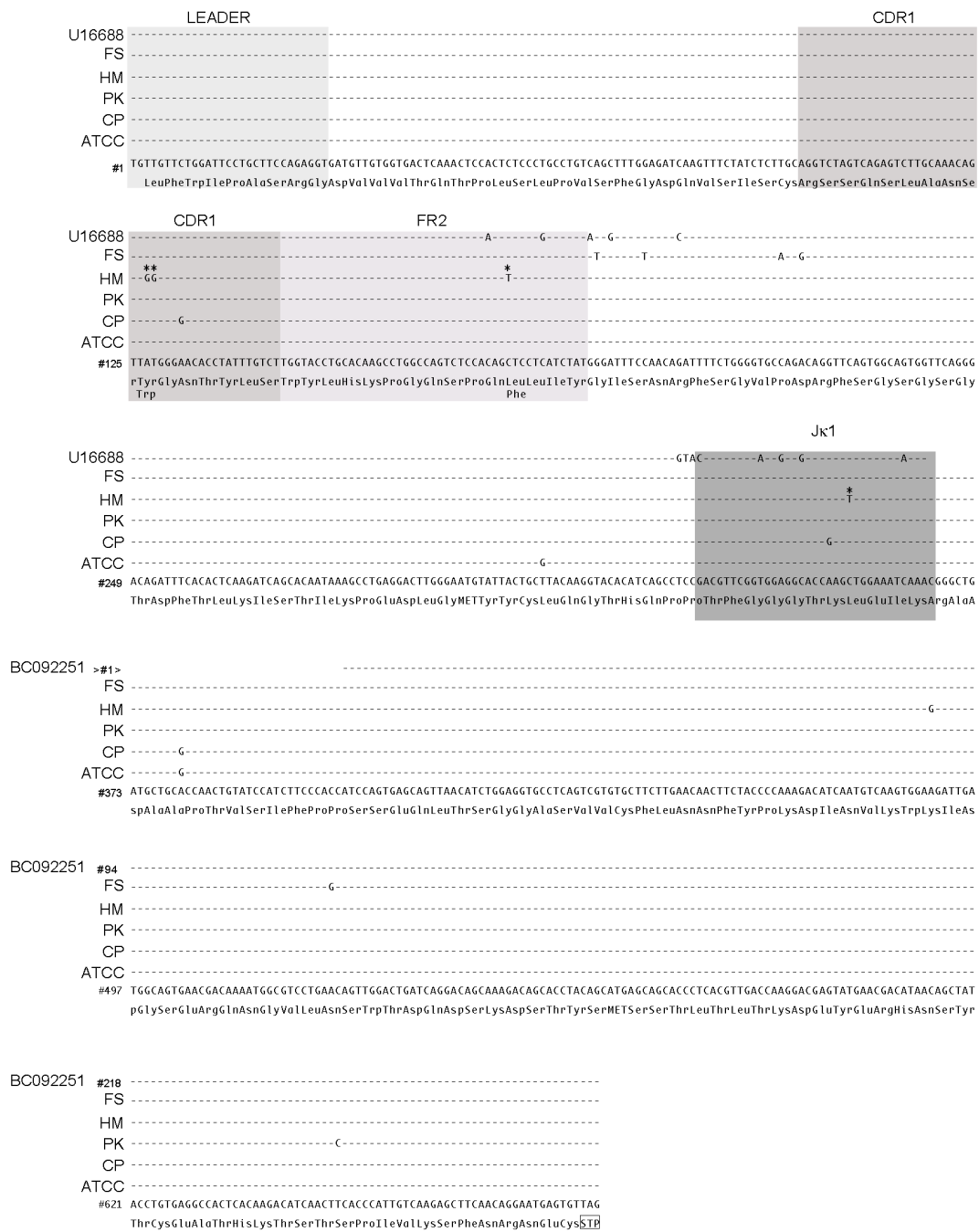
Fig. 20. κ chain cDNA comparison of WEHI-231 lines. Each line represents the sequence alignments of 6 sublones except for HM, with 5 subclones. Mutations that exist in all subclones are labeled (*). Reference sequences for the variable region: U16688 (BALB/c); and for the constant region: BC092251 (FVB/N). Leader sequence, CDR1, FR2 and J region are labeled in grey. Stop codon is marked by box.

From the comparison of the κ sequences we conclude, that from all WEHI-231 lines, that only the subclone HM for sure has several mutations in its Vκ. Because these mutations showed in all five bacterial subclones, we exclude the possibility that they were method-induced. Further, because the mutations result in three amino acid changes, we think that they may influence the protein structure of the L chain, in a

way that would disable the formation of intact Ig molecules. Since HM was selected for low surface IgM expression, the mutated Vκ may be the reason for it.

An exogenous substrate is hypermutated at a much lower rate than 18-81

From the limited sequencing we have done from the endogenous Ig locus, we cannot determine a mutation rate. However, the mutations we found were unselected, and therefore are likely to be the consequence of a mutator activity. In fibroblasts (Yoshikawa et al., 2002) and in Chinese hamster ovary cells (Martin and Scharff, 2002b), expression of AID is both necessary and sufficient to cause an exogenous substrate to hypermutate. We tested whether the FS line would also hypermutate an exogenous substrate. Our reporter gene consisted of an inactive EGFP that fluoresces upon reversion of a premature amber stop codon (TAG). This stop codon is embedded in the canonical RGYW motif, the target sequence of AID mediated mutagenesis (Wang et al., 2004a). Any single point mutation at the codon activates the reporter—except the unmodified transition from G to A, which only creates the stop codon TAA. We transduced the FS and 18-81 lines with the retrovirus-based construct that contained the reporter and a gene encoding resistance to puromycin, with the cells being kept continuously under puromycin selection. As a control, the reporter gene contained functional GFP, without stop codon. After three days, we performed flow cytometry on the lines (Fig. 21). The profiles were different for the two cell lines with the GFP containing the stop codon: in the 18-81 line, two mutant cell populations accumulated, one with high fluorescence intensity, and one with an intermediate intensity; in the FS line, there was only the intermediate intensity mutant population. We know that the high intensity GFP cells are generated by AID mediated mutagenesis, and we confirmed this by sequencing (not shown) (Klasen et al., 2005; Wang et al., 2004a). In this population, reversions of the stop codon TAG is usually accomplished by a G to C, and, less frequently, by a G to T transversion (Klasen et al., 2005; Meyer et al., 1986; Wang et al., 2004a).
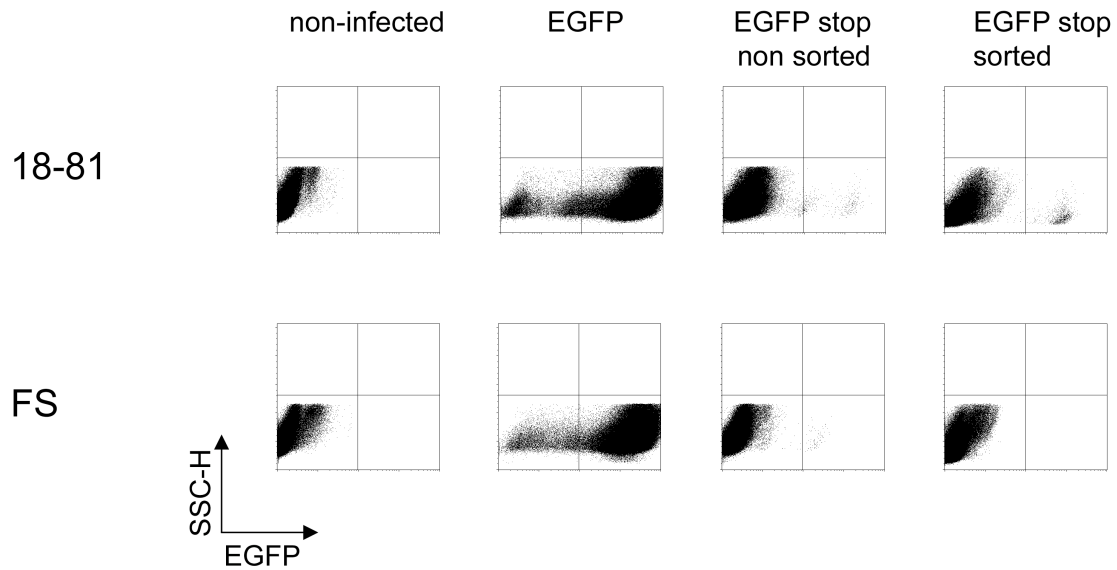
Fig. 21. Flow cytometry profiles of cells transduced with an exogenous indicator substrate for hypermutation 13 days after infection. Y-axis, side scatter; x-axis, fluorescence intensity, on a logarithmic scale, of enhanced GFP (EGFP). First column, 18-81 and FS cells, not transduced; second column, transduced with the indicator substrate containing an EGFP gene without a premature stop codon; third column, with premature stop codon; fourth column, sorted EGFP negative cell population, with premature stop codon, and expanded for 11 days.

We assume that these revertants in the population of intermediate intensity are generated by the reverse transcriptase of the retrovirus. We have three reasons for why we think that this assumption is correct: (i) we found such a population also in the 70Z cell line, where no AID is expressed. (ii) the intermediate fluorescence intensity of the revertants indicated to us that the stop codon had not reverted at the G nucleotide. Indeed, when we sequenced the revertants we found that the TAG invariably was converted into a TGG; the MLV reverse transcriptase has a mutation preference from A to G (not shown). Apparently, the Trp in this position, which is encoded by TGG, results in less than optimal fluorescence of GFP. Because of AID acting on G·C, the mutation expected to be generated is TAT or TAC, which encode tyrosine. (iii) Because the reverse transcriptase is present only at the time of transduction, such revertants do not accumulate; obviously dependent on the error frequency of the reverse transcriptase, they are generated only when a sufficient numbers of cells are transduced. Indeed, when we sorted the GFP negative population of 18-81 and FS, respectively, the revertant population of intermediate fluorescence intensity did not return after ten days in either cell line (Fig. 21). However, the high intensity revertant population returned in the 18-81 cell line, but not in the FS line. Fig. 22 shows a time course of accumulation of mutations in the reporter construct, as measured by flow cytometry. The duplicate cultures of FS,

designated A and B, did not accumulate any mutations over the time period assayed here, while the 18-81 line did. As it had been found, if the mutation rate is normalized by chronological time (Wang and Wabl, 2005a), the 18-81 line mutates the indicator plasmid at a rate of 10^{-5} per bp per day, while the FS line shows the much lower rate of 2.6×10^{-7} per day (Wang and Wabl, 2005b). This represents a 37-fold lower mutation rate in FS than in 18-81 (Wang and Wabl, 2005b), although the cDNA expression level of AID is 3.5 times higher in FS than 18-81 (Fig. 9). These results by Wang and Wabl (Wang and Wabl, 2005b) indicate that the rate in FS may be still enough to qualify for hypermutation.
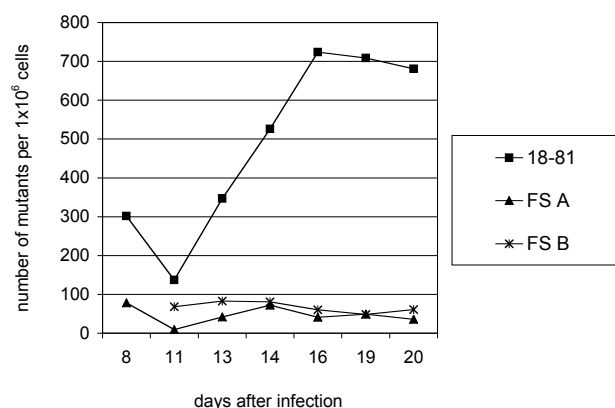
Fig. 22. Time course of accumulation of mutations in the EGFP gene with a premature stop codon. Y-axis, number of revertants per one million cells; x-axis, days. Squares, 18-81 cells; triangles and stars, two independent cultures of FS cells.

Hypermutation in precursor and progeny WEHI-231 lines

In order to test whether the putative 18-kD AID variant in the precursor cell lines is functional, we studied the generation of somatic mutations with our indicator plasmid. We now wanted to measure again the mutation frequencies, but also include the precursors PK and CP, and the FS subclone HM in the analysis. The WEHI-231 cell lines were transduced with the indicator plasmid, and after three days, cells were sorted against GFP positives, to purge false positives due to the error rate of viral reverse transcriptase. For these experiments, it is important to note that the transduction efficiencies were similar in the WEHI-231 lines and that same cell numbers were transduced.
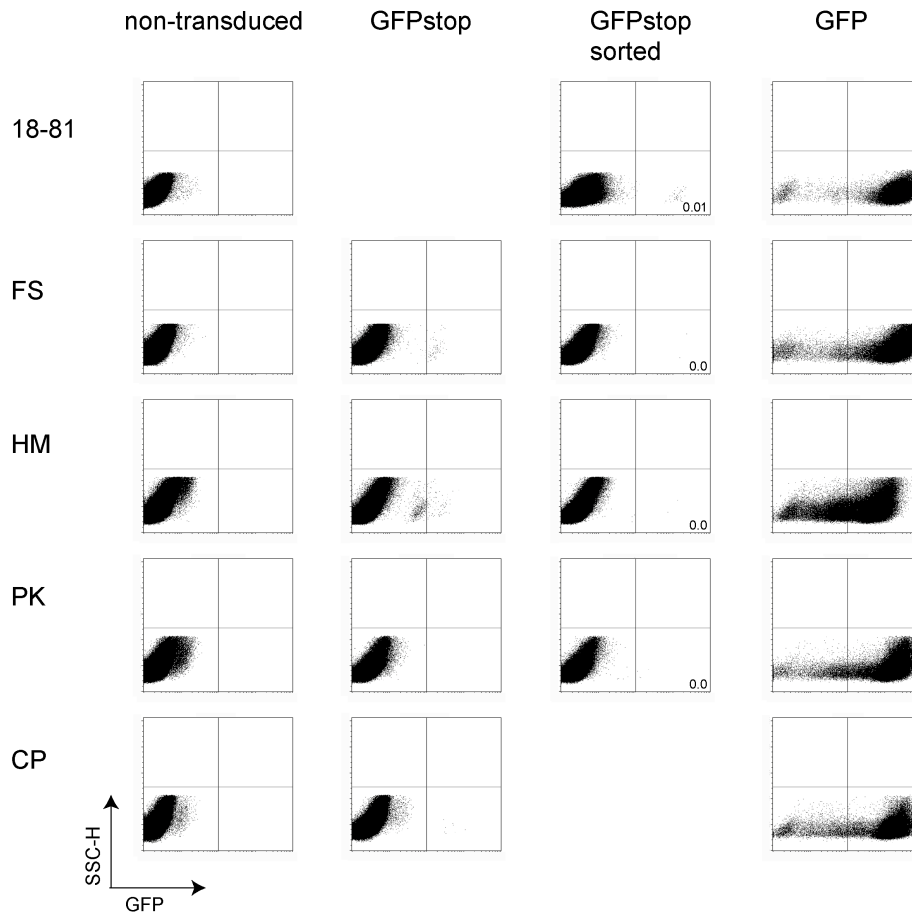
Fig. 23: Flow cytometry analysis of various WEHI-231 cell lines (indicated to the left of blots) 17 days after transduction with the indicator plasmid. GFPstop, cells transduced with indicator construct; GFPstop sorted, cells transduced with indicator construct and sorted against GFP-expressing cells three days after transduction; GFP, transduced with functional GFP. X-axis, green fluorescence intensity on a logarithmic scale; Y-axis, side scatter.

In this analysis (Fig. 23), we made the following observations: First, there are three kinds of mutant populations: one that is close to the right of the quadrant partitions (predominant in FS), and another that is close to the left of the quadrant partition (predominant in HM); and a third one that is of highest intensity (0.01% of the cells of GFP stop sorted in 18-81). From previous sequencing, we know that this high-intensity population is the consequence of AID activity, which reverts TAG into TAC or TAT. Such revertants are absent in all WEHI-231 lines, whether or not previously sorted, which is in line with the low mutation rate noted before.

The stop codon reversion of both intermediate populations are most likely due to the reverse transcriptase of the virus because they are absent in cells sorted after transduction. To identify the nature of the mutations, we sorted both populations from FS and HM and sequenced GFP from genomic DNA. We found that the intermediate GFP population from FS had reverted the TAG stop codon to TTG (encoding leucine). In HM, the intermediate GFP population reverted to TGG. An A to G

mutation is, indeed, the signature of the MLV reverse transcriptase. Whether or not the A to T mutations in FS have any other significance is not yet clear.

Lack of class switching

Because of its high-level expression of AID, it was possible that the FS line switches its IgM to other subclasses. However, isotypes other than μ have not been reported in WEHI-231. Given that the FS line has been in cell culture an undefined time, one would assume that if the line is actively switching, then other isotypes would accumulate. When we stained FS for all Ig subclasses, we found no evidence of switched cells (Fig. 24). In our flow cytometry experiments, we used commercially available antibodies to all eight classes. As positive controls we stained various hybridoma cell lines, which display very little Ig on their surface, because they secrete their Igs. However, putative switched cells of FS would be expected to show much higher intensity. We found that besides cells with IgM, in FS there are no cells that express any other Ig isotype (Fig. 24).
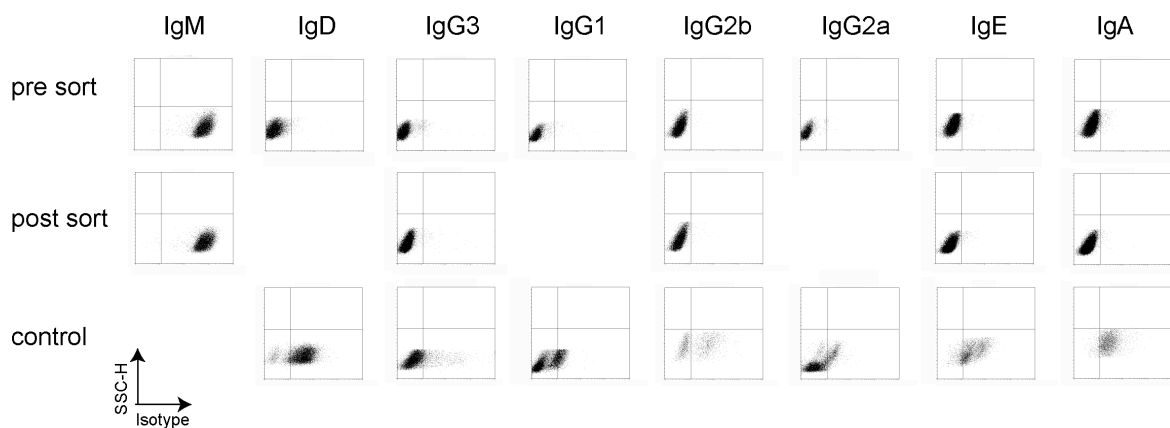
Fig. 24. Absence of Ig class switched cells in the FS line. (A) Flow cytometry assessing membrane expression of the eight H chain isotypes. Y-axis, side scatter; x-axis, fluorescence intensity on a logarithmic scale. The isotypes assayed are given above the respective flow profile. Pre sort, cells out of culture; post sort, fluorescence positive cells sorted and grown in cell culture. Control, hybridomas synthesizing the respective isotype.

In flow cytometry profiles, there are always some cells with high intensities, where none are expected. Thus we found a few such cells in our stainings for IgG1, IgG2b, IgE, and IgA. To decide whether these cells represent background, or rare switched cells, we sorted them and grew them in cell culture, to be flow profiled again (Fig. 24). The expanded cultures contained only cells with surface IgM, thereby confirming that FS does not switch its IgM class.

We also tried whether we could induce CSR in the FS cell line. Non-activated B cells stimulated with IL-4 and LPS switch their Ig class preferentially to IgG1 and, to a much lesser extend, to IgE. However, even after stimulation with IL-4 and LPS for 65 hr, the FS line expressed no other Ig class than IgM on its surface (Fig. 25). As a positive control, spleen cells were stimulated the same way. Nevertheless, the percentage for the IgE positive cells is too high and may be an artifact.
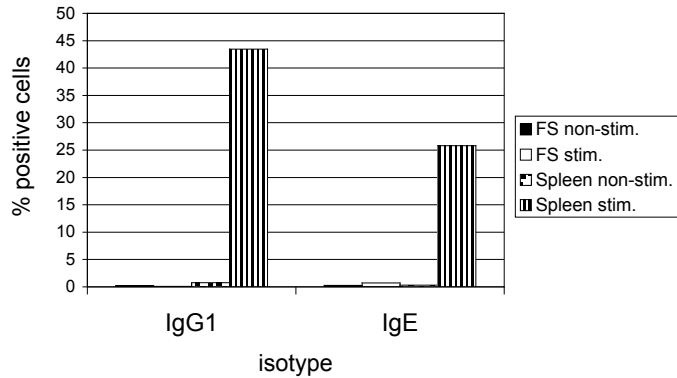
Fig. 25. Bar graph representation of isotype switching upon addition of IL4/LPS. Y-axis, percentage of isotype positive cells; x-axis, IgG1 and IgE isotypes, in non-stimulated and stimulated FS and spleen cells, respectively.

Because we were particularly interested whether FS would contain at least some IgE cells, we used PCR amplification of DNA between V and Cε primers, which amplify a functional ε gene.
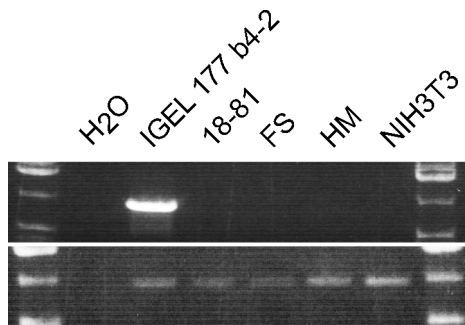
Fig. 26. RT-PCR amplification of the ε gene in FS and in control cell lines. The GAPDH gene served as a method and loading control.

Although we could easily amplify the ε gene from the IgE producing hybridoma IGELb4-2 (Rudolph et al., 1981), there were no bands in FS and its subclone HM, nor in the 18-81 line, which switches from μ to γ2b, or in the NIH3T3 line with no Ig genes (Fig. 26).

CSR indicator plasmid

The continuous expression of IgM is essential for the growth of WEHI-231 in the cell culture (Jakway et al., 1986; Page and DeFranco, 1988). It is not clear whether this requirement is strict in that it has to be Ig of the IgM class, or whether other classes would also support growth. As a consequence, it is also not clear whether upon CSR, when WEHI-231 would express an Ig other than IgM, the switched cell would apoptose due to the lack of IgM. To avoid potential apoptosis while testing for CSR, we transfected FS with an exogenous CSR indicator plasmid, denoted SCI(μ,α) (Okazaki et al., 2002) (Fig. 27). This construct contains the Sμ and Sα regions, which are directed by elongation factor 1α promoter (EF-1α) and a tetracycline-responsive promoter (pTET), respectively, because in order for switch recombination to occur, the S regions must be transcribed. The tetracycline-responsive promoter drives the transcription of the Sα, and is active in absence of tetracycline. Therefore, in order to test whether recombination is S region–mediated, tetracycline can be added and no recombination ought to occur. Both S sequences are removed by splicing from each transcript. The non-recombined S regions prevent the indicator CD8α from being expressed on the cell surface. This is because the coding sequences for the extracellular (EC) domain of CD8α, and the sequence for the transmembrane (TM) domain of CD8α fused with GFP, are separated into two transcription units. The extracellular domain of CD8α can be anchored on the cell surface only after its fusion with the transmembrane domain, which is enabled by recombination between the two S regions. The detection of CD8α by flow cytometry thus is a readout for switch recombination.
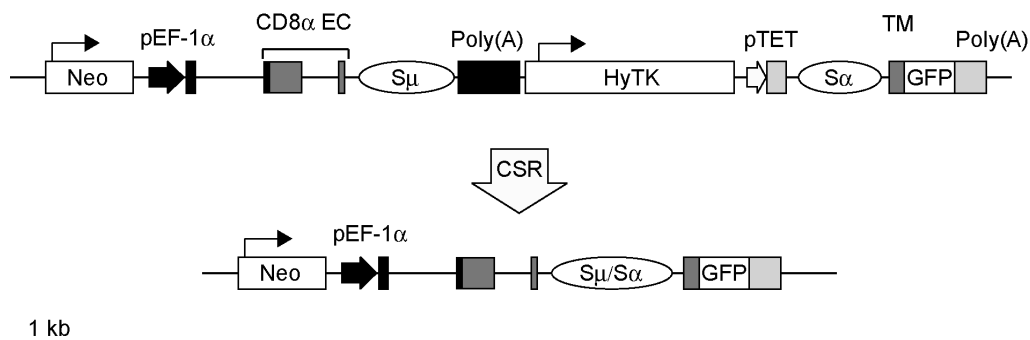
Fig. 27. Structure of the CSR reporter construct SCI(μ,α). pEF-1α, elongation factor 1α promoter; pTET, tetracycline-responsive promoter; Neo, neomycin resistance gene with the TK promoter; HyTK, gene for the fusion protein hygromycin phosphotransferase-thymidine kinase, with the CMV promoter; EC and TM indicate extracellular and transmembrane domains of CD8α, respectively. Upon CSR, S regions recombine, and CD8α and GFP are expressed.

To test the SCI(μ,α) we used 18-81 as a positive control, 70Z as a negative control, and FS as the cell line to test. First, we stably transduced these lines with a retroviral vector encoding a tetracycline-controlled transactivator pRevTet-Off. We then transiently transfected the cells with SCI(μ,α) and cultured them in the absence or presence of Doxycycline for 7 days. For flow analysis, cells were stained with an anti-CD8α antibody. As shown in Fig. 28, the SCI(μ,α) indicator switch construct did not work satisfactorily in our cell lines: we only saw very few switched cells (0.41 %) in our positive control cell line 18-81, although in this line the switch rate at the endogenous locus is 10^{-2} per cell generation. Also, we could not completely silence transcription from the pTET promoter, either because the doxycyline dose was too low, or the promoter is leaky. However, because we did not find GFP/CD8α double positive switched cells in FS or 70Z, we take this result as a weak, but an additional evidence for the absence of CSR in FS.
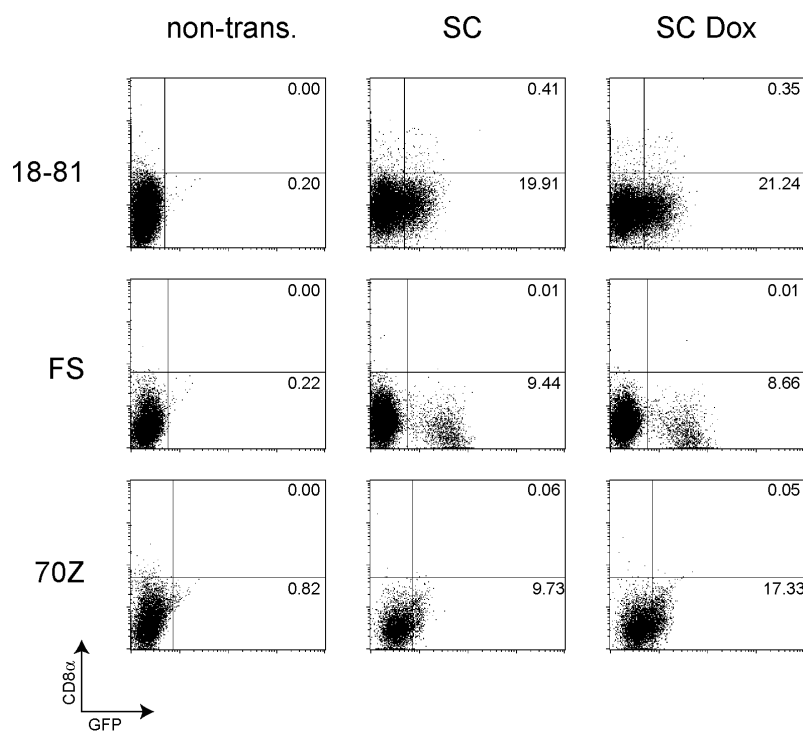
Fig. 28. SCI(μ,α) transfected FS, 18-81 (positive control), and 70Z (negative control). Transfected cells were cultured without (middle column) or with Doxycycline (right column). Non-transfected cells (left column). Doxycline is supposed to inhibit CSR by blocking transcription of the Sα region.

Germline transcripts

One can think of various reasons of why there are no cells with switched Ig isotype in FS. A reason for why the cells do not switch would be the lack of an open configuration of the locus the cell is going to switch to. Such an open configuration is indicated by the transcription of the so-called germline transcripts, which originate at a promoter upstream of the respective I exon, extend through the S region and C

region, and are spliced so that the I exon is joined to the C region (Gerondakis, 1990; Lutzker and Alt, 1988; Stavnezer-Nordgren and Sirlin, 1986). Such germline transcripts are harbingers of CSR (Stavnezer et al., 1988), and their splicing is necessary in order for switch recombination to occur (Lorenz et al., 1995). We, therefore, examined whether or not FS produces germline transcripts of any of the six constant region loci the cells might switch to. We RT-PCR amplified transcripts of $\gamma 2b$, ε and α, and found that FS expresses all of these, while the 18-81 cell line lacks the ε transcript and has a smattering of an α transcript (Fig. 29), and NIH3T3 has none of them.
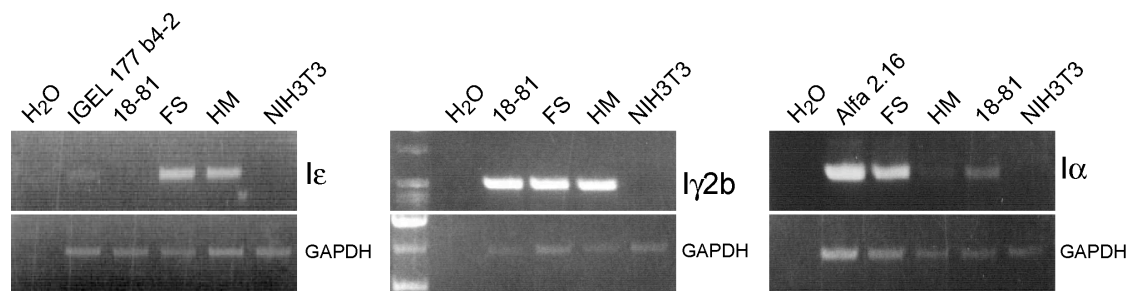
Fig. 29. Production of Ig germline transcripts in FS. (A) RT-PCR amplification of Iε, I$\gamma 2b$, and Iα transcripts. H$_2$O, amplification without template, but all other reagents.

Because class switching can occur on one allele only, or on both alleles, hybridomas that are derived from switched cells may or may not contain a germline transcript. Upon switching an allele, the gene encoding the germline transcript is lost on that allele. Apparently, the ALFA hybridoma, which was derived from a fusion between the I29 cell line and the plasmacytoma Ag8.653, retained the silent allele in the germline configuration, and, therefore, still synthesizes the α germline transcript. In the IGEL b4-2 hybridoma line, there is very little Iε transcript. We think that due to chromosome loss, there are few cells left with the silent homolog of chromosome 12, where the IgH is located. At any rate, we verified the identity of all these transcripts by sequencing.

The FS line seems to express as much $\gamma 2b$ germline transcript as the switching line 18-81, and almost as much a transcript as the ALFA hybridoma. But the quantitative outcome of RT-PCR can be very deceptive, and it is possible that the amount of transcript is much lower in FS and simply not enough to start the recombination process. Indeed, when we performed Taqman analysis for the Iε and I$\gamma 2b$ transcripts, this proved to be the case (Fig. 30). In comparison to the stimulated spleen cells, expression of the ε transcript was 74 times lower in FS and 195 times lower in the subclone HM (Fig. 30A, C). Expression of the $\gamma 2b$ transcript was 151 times lower in

FS and 370 times lower in the subclone HM than the cell line 18-81 (Fig. 30B, D). Thus, while it is not known how much transcription off the I promoter is needed in order to instigate the switch recombination process, there might not enough germline transcript expressed to enable FS to switch.
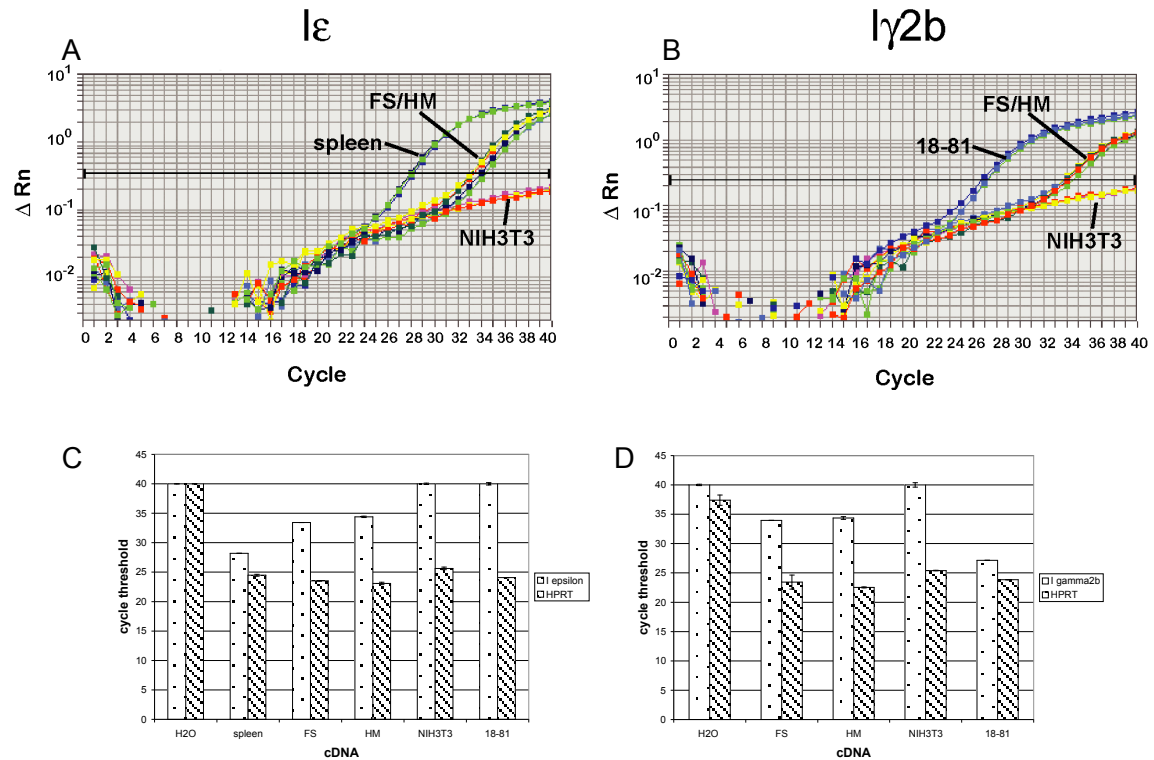

Fig. 30. Taqman analysis of (A) Iε transcript (B); and of Iγ2b transcript; (C) and (D) bar graph representations of the relative expression levels of Iε and Iγ2b transcripts in the various cell lines. X-axis in A and B, number of cycles needed to reach a preset value, on the y-axis; i. e., the lower the number of cycles, to higher the expression level of the cDNA. HPRT, cDNA encoding the enzyme hypoxanthine guanine phosphoribosyl transferase.

There remained the possibility that the gene segments downstream of Cμ might have been deleted on the active homolog of the FS cell line. In that case, no switching to any of the other classes would be possible. To address this question, we analyzed the γ2b, γ2a, and α loci by PCR amplification and sequencing.

Because WEHI-231, and thus FS, is of (BALB/c x NZB)F$_1$ origin, it has two different Ig haplotypes. The haplotype derived from BALB/c is Ig[a], and fully sequenced, whereas the NZB haplotype is less defined. However, the IgM expressed on the surface of this line reacts with an antibody to the b allotype (Fig. 24, 32), i.e., non-BALB/c; and by sequencing, we could assess it as being non-Ig[a] as well.

```
A   Ref. X16997  --------------T-----G---------------------------------------------C-------------------T-------
    FS I FOR.    --------------T-----G---------------------------------------------C-------------------T-------
    FS I REV.    --------------T-----G---------------------------------------------C-------------------T-------
    FS II FOR.   --------------C-----A---------------------------------------------T-------------------A-------
    FS II REV.   --------------C-----A---------------------------------------------T-------------------A-------
    Ref. J00470  --------------C-----A---------------------------------------------T-------------------A-------
                 ATGATCTCCCTGAGYCCCATRGTCACATGTGTGGTGGTGGATGTGAGCGAGGATGACCCAGATGTCCAGATCAGCTGGTTYGTGAACAACGTGGAAGTACWCACAGCT
                 METIleSerLeuSerProMETValThrCysValValValAspValSerGluAspProAspValGlnIleSerTrpPheValAsnAsnValGluValLeuThrAla
                      Ile                                                                                        His


B   Ref. J00461  ----------------------------------------------------------------------------------------------------
    FS III FOR.  -A---------TT---------------------------G----------------------GC----TA-----A-------------------
    FS III REV.  -A---------TT---------------------------G----------------------GC----TA-----A-------------------
    FS IV FOR.   ----------------------------------------------------------------------------------------------------
    FS IV REV.   ----------------------------------------------------------------------------------------------------
    HM I FOR.    -A---------TT---------------------------G----------------------GC----TA-----A-------------------
    HM I REV.    -A---------TT---------------------------G----------------------GC----TA-----A-------------------
    HM II FOR.   ----------------------------------------------------------------------------------------------------
    HM II REV.   ----------------------------------------------------------------------------------------------------
                 CTCCTAACCTCGAGGGTGGACCATCCGTCTTCATCTTCCCTCCAAATATCAAGGATGTACTCATGATCTCCCTGACACCCAAGGTCACGTGTGTGGTGGTGGATGTGA
                  ProAsnLeuGluGlyGlyProSerValPheIlePheProProAsnIleLysAspValLeuMETIleSerLeuThrProLysValThrCysValValValAspVal
                    Leu                              Lys                      Ser   Ile


C   IgA a allele ---C--TC----GC-----T--C-----------------------------------T--------------------------------------------
    IgA b allele ---G------------------------------------------------------------------------------------------------
    IgA d allele ---G------------------------A-----------------------------------------------------------------------
    IgA f allele ---G------------------------------------------------------------------------------------------------
    FS FOR.      ---C------------------------------------------------------------------------------------------------
    HM FOR.      ---C------------------------------------------------------------------------------------------------
                 CACSAGCTCTGTCAAGTGACCCAGTGATAATCGGCTGCCTGATTCACGATTACTTCCCTTCCGGCACGATGAATGTGACCTGGGGAAAGAGTGGGAAGGATATAACCACC
                  ArgAlaLeuSerSerAspProValIleIleGlyCysLeuIleHisAspTyrPheProSerGlyThrMETAsnValThrTrpGlyLysSerGlyLysAspIleThrThr
                    ProVal   Cys                                         Phe
```

Fig. 31. Partial sequences of the γ2a, γ2b and α alleles. (A) Alleles of Cγ2a in FS (sequences I and II), sequenced in forward (FOR.) and reverse (REV.) orientation. Ref. X16997, Genbank sequence of Ig[b]; Ref. J00470, Genbank sequence for Ig[a]. Underneath the consensus sequence, translation of the b allele into polypeptide; and amino acid residue substitutions between the alleles. (B) Alleles of Cγ2b in FS (sequences III and IV), and subclone HM (sequences I and II), sequenced in forward (FOR.) and reverse (REV.) orientation. Ref. J00461, Genbank sequence for Ig[a]. (C) a alleles of various allotypes and the non-Ig[a] alleles in FS and HM, sequenced in forward (FOR.) orientation. Translation as from the b allele.

In some instances, sequences deposited in the Genbank exactly matched allelic fragments of non-Ig[a] allotypes. At least exons 1 to 4 of γ2b and γ2a were present on both alleles the FS cell line, as well as some sequences of the active a allele (non-Ig[a]) (Fig. 31). While this does not exclude defects in the switch regions, or in sequences not covered, they would have to be present in the γ2b, γ2a and α loci; but certainly, there is no wall-to-wall deletion downstream of Cμ.

FS subclone HM

This chapter focuses on the FS subclone HM, which was selected for decreased surface IgM expression.

IgM allotypes in HM and FS

Expression of surface IgM is crucial for survival and growth of the cell line WEHI-231. Because WEHI-231 is derived from the F_1 mouse of BALB/c and NZB, it has one allele of the a allotype (BALB/c) and one of the n allotype (NZB), which is similar to the b allotype (C57BL/6). The allotypic difference of the IgM of a allotype (IgM^a) and IgM of b allotype (IgM^b) is detectable by monoclonal antibodies (Nishikawa et al., 1986) and is encoded at position 292 in the first exon of the C region of μ (Cμ1). At this position the b allotype has an adenine encoding lysine (codon AAA), while the a allotype has a guanine encoding arginine (codon AGA) (Schreier et al., 1986). When we stained for IgM of a and b allotypes, we found FS to be positive for IgM^b and negative for IgM^a. In contrast HM is positive for IgM^a and negative for IgM^b (Fig. 32A). Moreover, the total IgM expression of HM is 10 times lower in comparison to FS, and the decreased surface μ chain expression correlates with surface κ chain in HM (Fig. 32B)
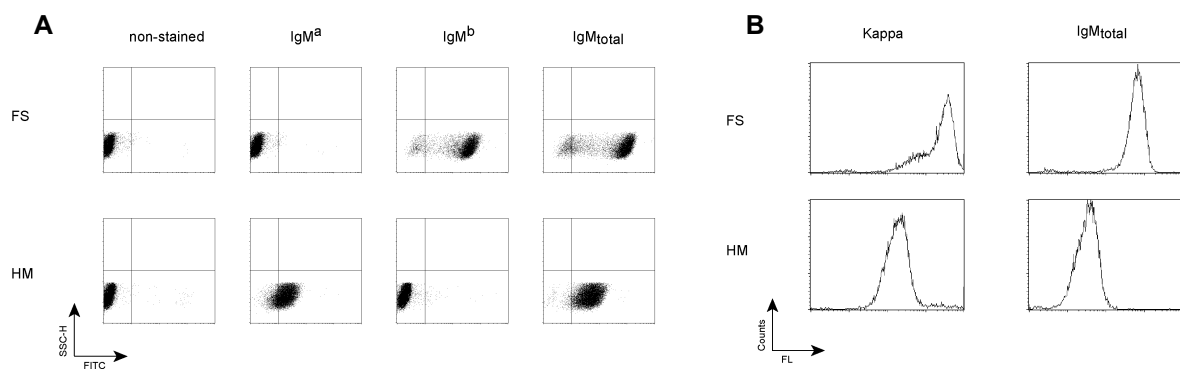
Fig. 32. (A) Cell line FS and its subclone HM stained for IgM^a, IgM^b and IgM^{total} expression. FS, upper, HM, lower row. X-axis, fluorescence intensity (logarithmic scale) of FITC coupled antibody, Y-axis, side scatter. (B) Histogram of κ and IgM^{total} expression in FS and HM. X-axis, fluorescence intensity (logarithmic scale) of FITC (anti-IgM^{total}) or PE (anti-κ) coupled antibody, Y-axis, number of events.

This switch in allotypes was puzzling. We thought the FS line, or the HM clone, has switched during culturing, and so we profiled cell lines from early freezings of WEHI-231 from ATCC (designated ATCC), and the labs of Christopher Paige (CP) and Paul Kincade (PK). These 'early' lines express IgM of a allotype, as shown by flow cytometry (Fig. 33). From this we hypothesized that the FS cell line, while cultured for

an unspecific time, has undergone one or multiple differential steps that lead to the switch to IgM^b. In addition, the switched founder cell in FS must have overgrown its precursors.
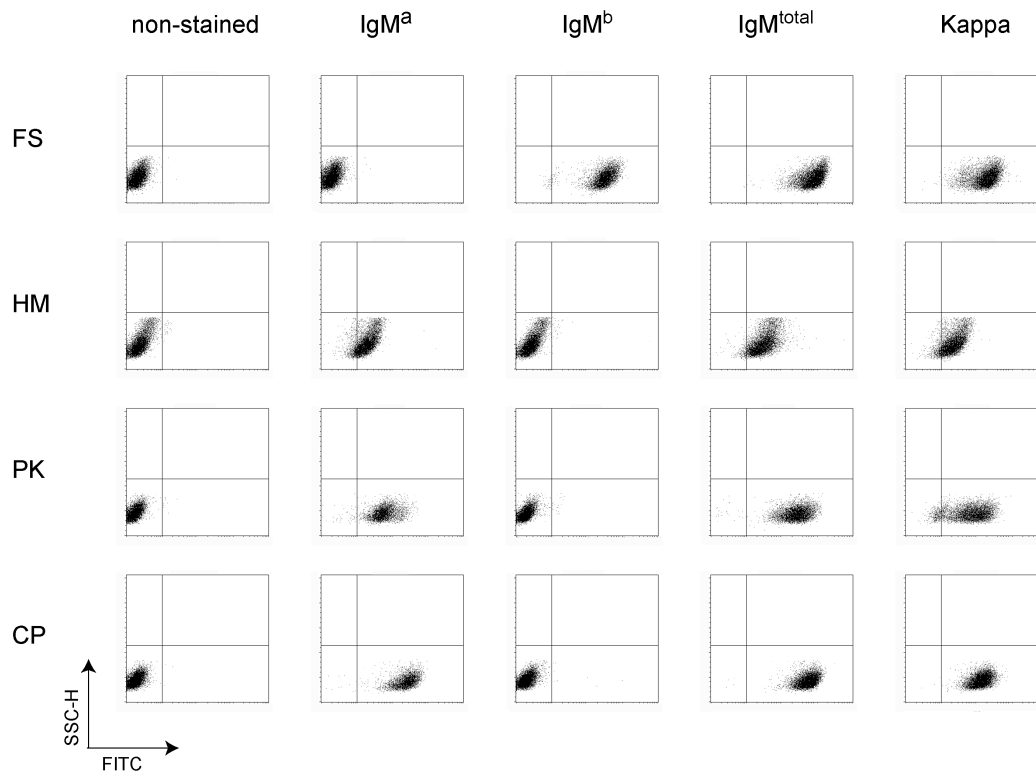
Fig. 33. Flow cytometry analysis of WEHI-231 lines FS, PK and CP and FS subclone HM stained with FITC coupled antibodies against IgM^a, IgM^b and kappa (monoclonal) and IgM^{total} (polyclonal). X-axis, fluorescence intensity (logarithmic scale) of FITC coupled antibody, Y-axis, side scatter.

To confirm that FS, HM, PK and CP originate from the same B lymphocyte we sequenced part of the cDNA that covers the VDJ exon linked to Cμ of these cell lines. The BALB/c and NZB alleles differ in four single nucleotides (SNPs). Fig. 34 shows the positions and identities of these SNPs for the b allele above the line and for the a allele below the line. The allelic difference furthest to the left translates into an amino acid residue difference that is epitope for the allele-specific monoclonal antibodies (see above). The remaining three allelic differences are located in introns. The NZB allele is not as well defined as the one of BALB/c (a allotype) or C57BL/6 (b allotype). However, when we sequenced the Cμ locus of the NZB mouse it showed great homology to the C57BL/6 sequence.
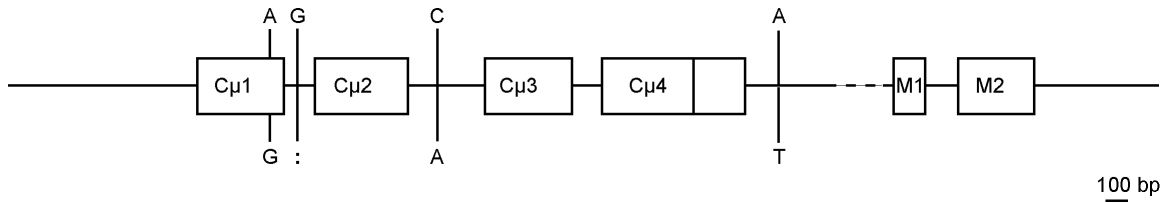
Fig. 34. Cµ locus with allelic differences between NZB (above line) and BALB/c (below the line) mice.

We performed RT-PCR with oligonucleotides priming specifically in the V region of the productive (V_P), as well as of the non-productive (V_{NP}) VDJ rearrangement, and in the second exon of the Cµ heavy chain (Cµ2). With this primer combination, we could not amplify the non-productive VDJ rearrangement of FS (Fig. 35). Furthermore, we amplified and sequenced cDNA with 5' primers specific for exon borders $V_P J_H 2$ and $V_{NP} J_H 3$ in combination with oligo dT primer in order to see what is linked to the productive allele. Again, no amplification product was detected for the silent allele in FS.
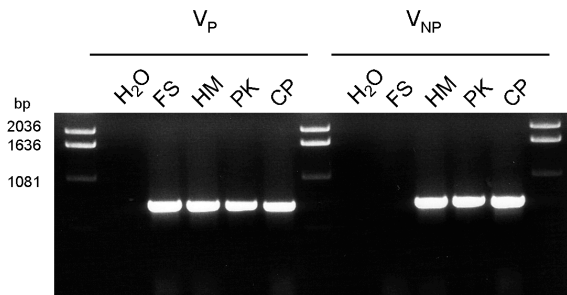
Fig. 35. RT-PCR of V-Cµ2 region from FS, HM, PK and CP. V_P: productive VDJ rearrangement; V_{NP}: non-productive VDJ rearrangement. Size of amplification product: ~300 bp.

However, when we amplified the V regions with primers for V_P and $J_H 2$ (productive VDJ rearrangement) and primers for V_{NP} and $J_H 3$ (non-productive VDJ rearrangement) on genomic DNA from FS and HM, we found the non-productive VDJ arrangement to be present in FS (Fig. 36).
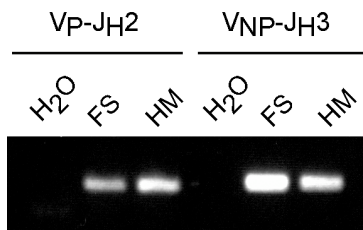
Fig. 36. Amplification of V-J region on genomic DNA extracted from FS and HM. V_P: productive VDJ rearrangement; V_{NP}: non-productive VDJ rearrangement. Size of amplification product: ~300 bp.

Subcloning in bacteria and sequencing of the V_P and V_{NP} of FS, HM, PK and CP, respectively, showed that in FS, V_P is linked to $C\mu^b$; in contrast we found in HM, PK and CP that V_P is linked to $C\mu^a$ and V_{NP} to $C\mu^b$ (Fig. 37).
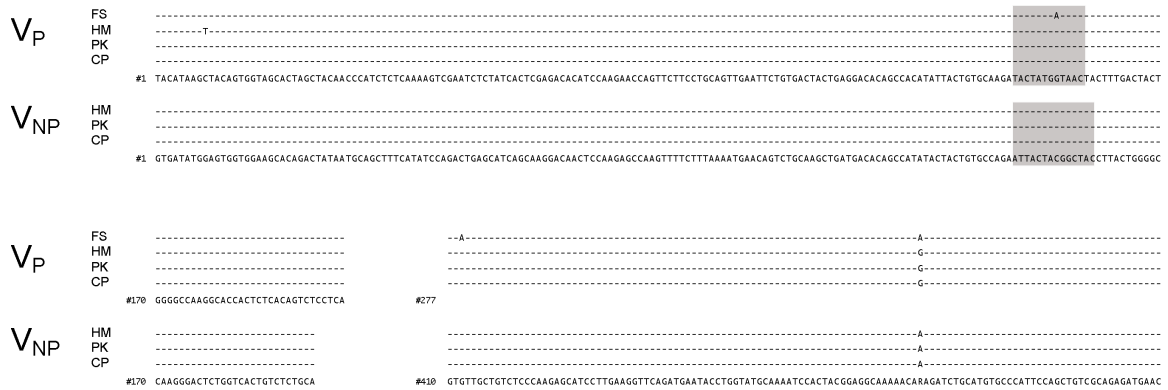
Fig. 37. cDNA encoding for either the productive VDJ (V_P) or the non-productive (V_{NP}) rearrangement linked to Cµ from FS, HM, PK and CP. First part of the sequence is V-, D- (grey area) and J-region. After the gap: Cµ1 with the allelic difference. Consensus sequence is V_P.

We amplified and subcloned (bacterial subclones) the genomic locus encoding Cµ from FS and HM and compared these sequences of BALB/c and NZB, respectively. Seven out of seventeen sequences from HM were of the b allele, and ten of the a allele. All thirteen sequences from FS were from the b allele. A summary for genomic loci encoding the Cµ heavy chain of FS and HM is schematically presented in Fig. 38. Upon allele switching in FS, the productive VDJ is linked to the $C\mu^b$ whereas $C\mu^a$ is lost. The subclone HM has both alleles: the productive allele is linked to $C\mu^a$, and the non-productive to $C\mu^b$. Based on the results shown in Fig. 37 we predict CP and PK to be the same genotype as HM and therefore included these WEHI-231 precursors in Fig. 38.
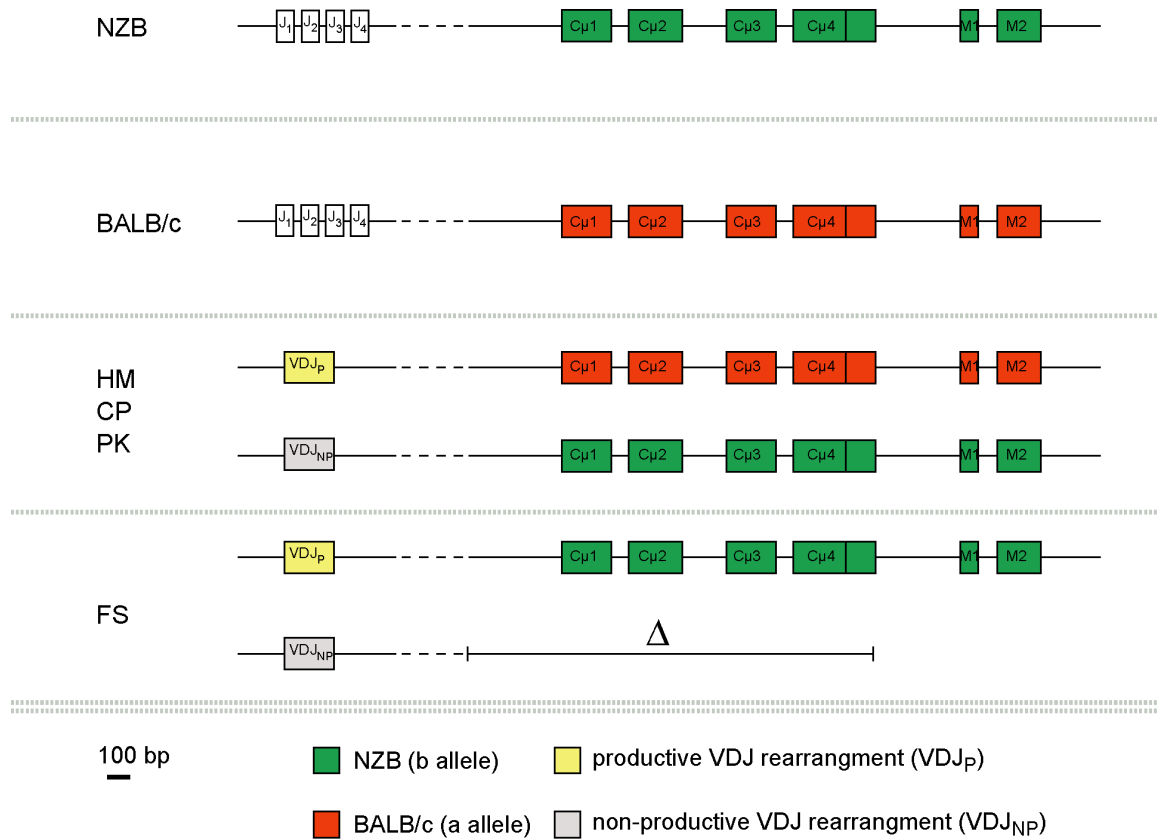
Fig. 38. Scheme of genomic loci encoding for the Cµ heavy chain of FS and HM, CP and PK in comparison to NZB and BALB/c. The a allele is labeled red; the b allele green; productive VDJ rearrangement yellow; non-productive VDJ rearrangement grey.

Based on these studies we conclude that in the HM precursor line (Fig. 6) at least one cell recombined its productive VDJ rearrangement with Cµ of b by switching the alleles and gave rise to the cell line FS. Because the HM precursor expressed AID, it is possible that such allele-switching is AID mediated.

Because we thought that the switch region for Cµ (Sµ) was involved in the process of allele switching, we tried to amplify Sµ. Because Sµ is highly repetitive over a range of ~3 kb, the amplification is challenging. For the germline sequence, the expected size of amplification product of the PCR shown in Fig. 39 is 3 kb. The detected amplification products, however, are only ~1 kb, and less, in both FS and HM.
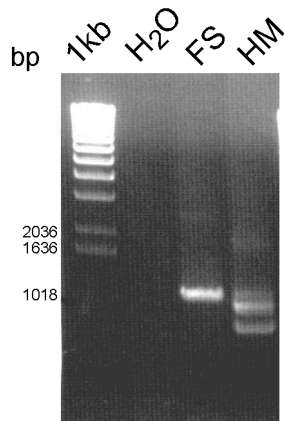
Fig. 39. Amplification of the Sµ from FS and HM.

Thus, barring an experimental artifact, recombination at Sµ has taken place in both FS and HM. By itself however, this does not mean that the Sµ region mediated the switch in alleles. Nevertheless, it is consistent with this notion, and more extensive experiments on the frequency of the process will allow us to decide whether or not to follow this line of experimentation. For sure the band pattern is different in FS compared to HM. So far, sequence analysis gave no clear result due to random annealing of the primers in this repetitive locus.

In summary, IgM of the a allotype is normally expressed in WEHI-231 lines. However, in the WEHI-231 descendant FS, the IgM allotype has been switched to the allotype of b. Presumably as a consequence of the switch, Cµ of the a allotype is deleted in FS, whereas there are still two alleles for the VDJ rearrangement. Because the Sµ regions of FS and the non-switched HM differ, we suspect that Sµ is involved in the switching of alleles. Due to the facts that FS expresses AID at a high level, and the involvement of the Sµ, we hypothesize, that the switching of alleles is AID mediated.

Decreased expression level of IgM in HM

As mentioned above, clone HM was selected for low expression of surface IgM. Figs. 32 and 33 show that HM expresses 10 fold less surface IgM than FS. The question arose what causes this decreased level. To address this, we first investigated the more trivial reasons: (i) switching to an Ig class other than IgM, (ii) secretion of IgM, rather than surface deposition (iii) retention of IgM within the cell.

The possibility that HM has switched to another subclass was partially ruled out by staining for IgG2b, IgE and IgA (Fig. 40), leaving open the possibility of a switch to IgG2a, IgG3 and IgG1. Very few cells appeared slightly positive. Those cells were sorted and taken in culture and analyzed by flow cytometry for a second time. They turned out to be false positive.
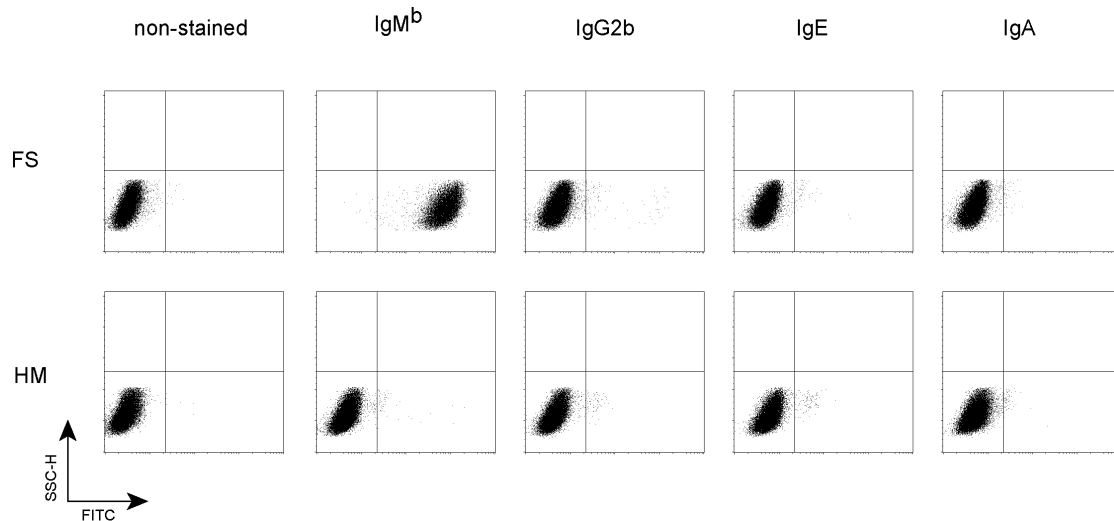
Fig. 40. Staining for divers Ig subclasses indicated above blots. Both FS (upper row) and HM (lower row) stained negative for subclasses IgG2b, IgE and IgA. X-axis, fluorescence intensity (logarithmic scale) of FITC coupled antibody, Y-axis, side scatter.

To address possibilities (ii) and (iii), that in subclone HM, IgM is secreted or blocked from reaching the surface, and accumulates in the cytoplasm, we performed an ELISA and immunofluorescence. We found some IgM in the supernatant of cultured FS, but none in the subclone HM, cultured under the same conditions (Fig. 41).
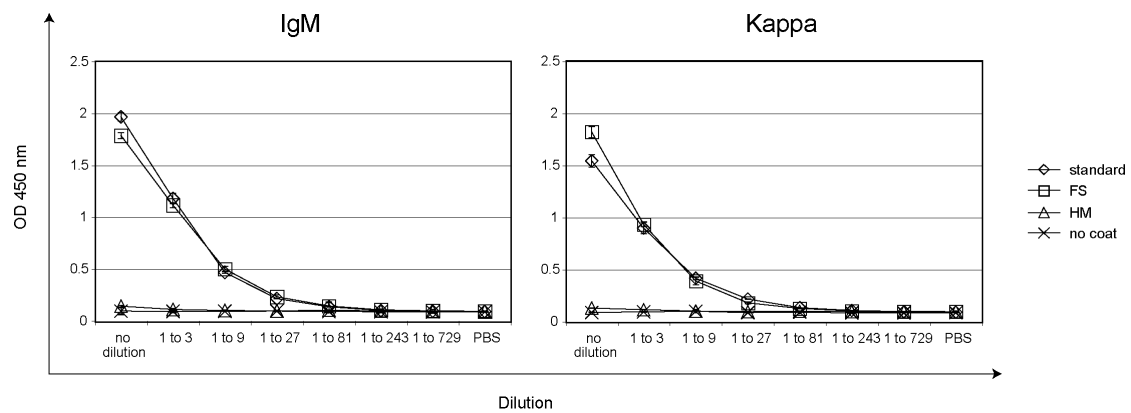
Fig. 41. Detection of secreted IgM by ELISA. Left graph: detection of μ chain; right graph detection of κ chain. X-axis: dilution of cell supernatant; Y-axis: OD at 450 nm. N=3

FS   HM
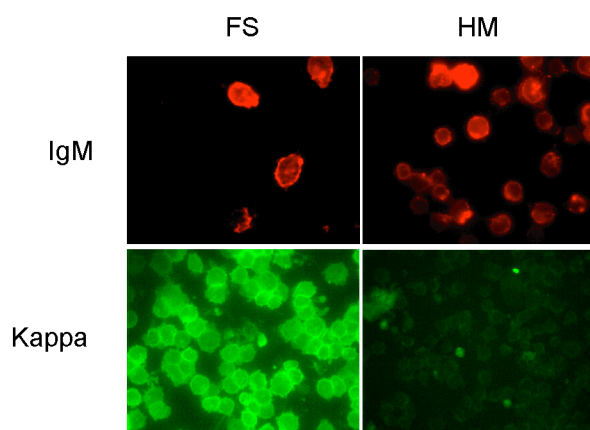
Fig. 42. Immunofluorescence with antibodies against μ chain (upper row) and κ chain (lower row) of FS (left) and HM (right).

In κ deficient cells, μ chains remain associated with the endoplasmic reticulum (ER) resident chaperone BiP in the ER lumen (Bornemann et al., 1995). In our immunofluorescence staining we found equal expression of μ expressed in FS and HM, but remarkably less expression of κ chain in HM (Fig. 42). Because of the decreased κ chain expression, we assume that the μ chain is retained in the lumen of the endoplasmic reticulum in HM. One reason for low κ protein may be the highly mutated Vκ of HM (Fig. 20), which gives rise to two amino acid changes. The amino acid exchanges may cause higher susceptibility to degradation and therefore loss of κ protein. In addition the amino acid changes may result in a conformational change of the κ protein, which make it incompatible to form with the μ chain a complete Ig molecule. Also, perhaps we cannot detect κ protein, because the anti-κ antibody epitope may be altered as a consequence of the mutations.

We as well determined the expression level of μ and κ by Western blot (Fig. 43) and confirmed the immunofluorescence data: The band for the κ chain is almost not visible in HM. In addition, there is a double band for IgM in FS, whereas there is only the lower band present in HM; it thus appears that HM has non-typical H and L chain expression (Fig. 43B). The μ, κ secreting hybridoma cell line SP1 was used as a positive control. However, although κ is strongly expressed, we could not detect a band for μ. When kept in culture for some time, hybridomas tend to lose H chain expression due to the loss of chromosome 12, and this seems to be the case here as well.

A

B
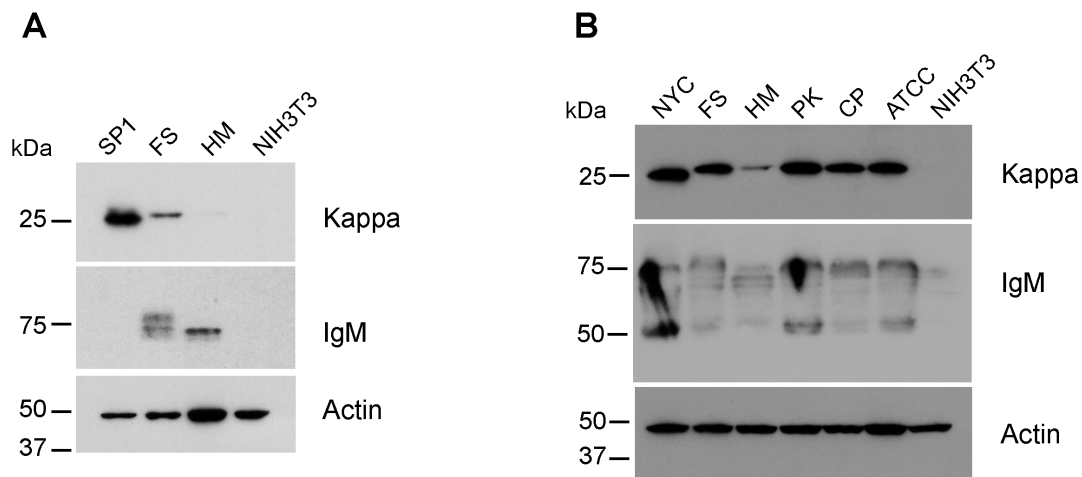
Fig. 43: Western blot analysis on κ chain and μ chain expression. (A) Comparison of HM with FS; SP1, positive control; NIH3T3, negative control. (B) Comparison of HM with all other WEHI-231 lines. NYC, positive control; NIH3T3 negative control.

An explanation for the missing upper IgM band in HM is defective glycosylation. Normally, IgM is glycosylated at six asparagines in the C region. In order to find out, we sequenced the whole Cμ of HM, but have not found any mutation at one of the glycosylated sites. Therefore, the lesion does not appear in the H chain itself. Consequently, the glycosylation defect must be in trans. To test whether or not there is defective glycosylation, we treated whole cell protein lysates of the WEHI-231 lines and subclone HM with the amidase N-PNGase F, an amidase, which cleaves the innermost GlcNAc and asparagine residues of high mannose, hybrid and complex oligosaccharides from N-linked glycoproteins (Maley et al., 1989). Indeed after treatment with N-PNGase F, bands from all WEHI-231 lines had the same size (Fig. 44). This shows that he μ chain of HM is underglycosylated. The question, though, remains whether or not the primary structure of μ is the cause for this.
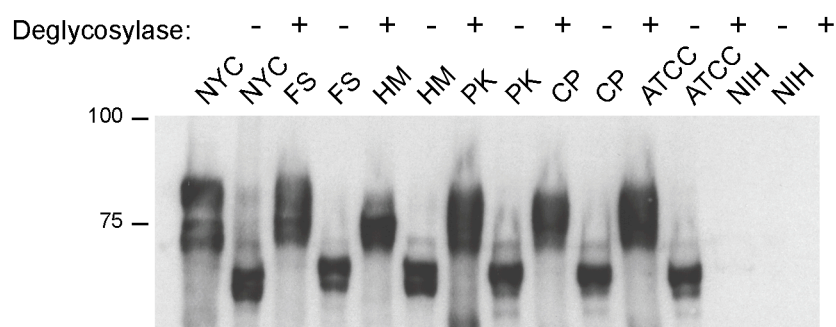
Fig. 44. IgM Western blot of samples treated with or without amidase N-PNGase F. IgM band of HM compares to other WEHI-231 after deglycosylation.

In summary, low surface IgM in HM is caused by low κ expression. Because we did not find mutations in the κpromoter, we consider the low κexpression a consequence of the mutated Vκ. Presumably, due to the mutations the κ chain is prone to degradation when formation of the complete Ig molecule with the H chain is perturbed. We further suggest that the defect to assemble IgM is responsible for the incomplete glycosylation of the μ chain in HM.

Retroviral insertional mutagenesis

As discussed before, the WEHI-231 line FS has the factors known to be necessary and presumed to be sufficient for class switch recombination and somatic hypermutation. Because we could not find any evidence that either mechanism efficiently works in this cell line, we assume that there must be additional factors that mediate the two processes. These factors might be activator or suppressors. In an attempt to identify (some of) the missing factors, we used retroviral insertional mutagenesis to switch them on or off. With this method, cells are infected with retrovirus, the provirus of which stably integrates into the host genome. In this way, genes can be disrupted, or turned on via the enhancer/promoter of the provirus. In order to estimate how many infected or transduced cells with insertion sites it takes, we do the following calculation: The haploid mouse genome is 3×10^6 kb. If the effect of the insertion extends over a length 10 kb, then we need at least 3×10^5 infected cells to cover the whole genome; this is under the assumption that the integration of the virus is at random sites.

AKV

A virus commonly used for the purpose of insertional mutagenesis is the AKV, and so we tried to use this virus first. Akv is a well-characterized ecotropic murine leukemia virus (MLV) (Buchhagen et al., 1980; Etzerodt et al., 1984; Jensen et al., 1986; Lovmand et al., 1990; Morrison et al., 1991; Nielsen et al., 1992); the cell surface receptor for Akv and other ecotropic MLVs is a cationic amino acid transporter, CAT-1 (Wang et al., 1991). For our experiments, we used a virus that has a GFP gene with an internal ribosomal entry site (IRES), which is within the 3'UTR (Wang et al., 2004b). After transfection of NIH3T3 with the virus-encoding vector, the virus-containing supernatant was added to the FS line.
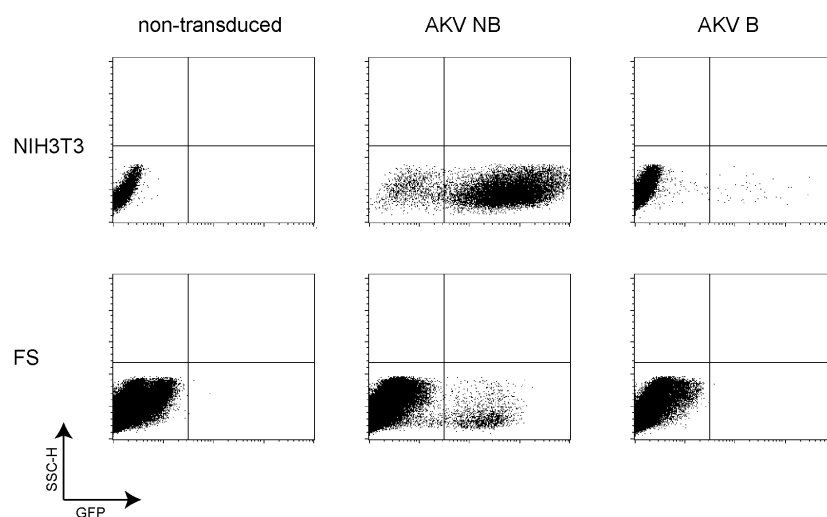
Fig. 45. Upper plots: NIH3T3 cells transfected with vectors encoding AKV NB and AKV B containing GFP as a marker after 9 days. Lower plots: Supernatant of NIH3T3 cells was applied to FS on day 20 after transfection. FS were analyzed 9 days after transduction. X-axis: GFP fluorescence on a logarithmic scale; Y-axis: side scatter.

Fig. 45, upper row, shows NIH3T3 cells transfected with vector encoding AKV of NB or B on day 9 after transfection. The NB and B types differ in their ability to infect cells of different genotypes. In our case, only 0.39 % cells are GFP positive with the B virus, whereas 96.65 % are GFP positive for the NB virus. For the purpose of insertional mutagenesis, the virus containing supernatant of the transfected NIH3T3 cells was added to the FS culture, 20 days after NIH3T3 transfection. Transduced FS cells were subsequently analyzed by flow cytometry 9 days after transduction. No GFP positive cell was found for FS cells infected with the B virus. Because NIH3T3 were not transfected in the first place, this was not much of a surprise. But also only 0.24 % of the cells appear to be GFP positive for the FS transduced with AKV NB. With such low infection rate, we would not cover the whole genome and therefore we decided to not use AKV for the purpose of retroviral insertional mutatgenesis.

Moloney virus

With AKV, we tried to use fully replication–competent virus. But for the purpose here, this replication competence is not needed as long as the transduction event is sufficiently frequent. With the Moloney-virus based mutation indicator plasmid, we hand in a retroviral construct that integrated into DNA with great efficiently. But it also had the additional benefit to immediately indicate potential mutator mutants that would arise from insertional mutagenesis. Initially, 3×10^6 FS cells were infected and selected with puromycin 24 hrs after infection. The infection efficiency was ~8 % (~4.8 x 10^5 infected cells). Cells were sorted against GFP positives 3 days after infection in order to purge the green fluorescent cells due to reverse transcriptase of

the virus. Cells were sorted and subcloned for GFP positives 29 and 52 days post infection, respectively. On day 87 days after infection, 5 subclones were screened by flow cytometry (Fig. 46).
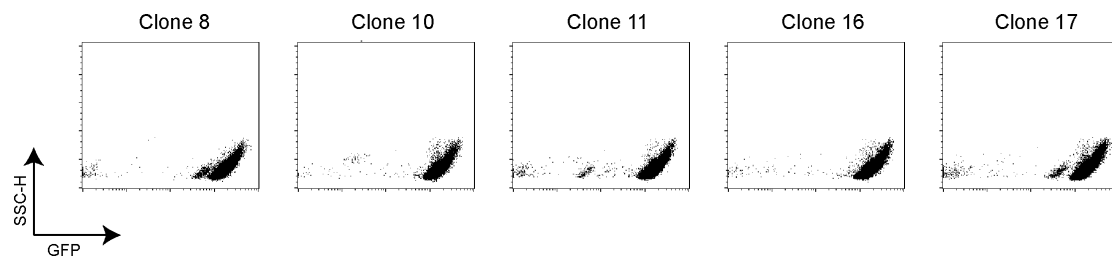
Fig. 46. FS transduced with the indicator plasmid containing a GFPstop was sorted and subcloned for GFP expressing cell. Five subclones were analyzed by flow cytometry. X-axis: GFP fluorescence on a logarithmic scale; Y-axis: side scatter.

The two subclones 11 and 17 showed in addition to the major high-fluorescent intensity population a second population with lower GFP expression. This was interesting, as we assumed that such populations may be the result of an active mutator, and the GFP would lose fluorescent intensity due to accumulation of mutations. If this would indeed be the case, then the GFP intensity should be lost over time. We then sorted four different populations of clone 11, 91 days after transfection (Fig. 47). We refer to these subpopulations as 11.1 (lowest GFP intensity) to 11.4 (highest GFP intensity). They were kept in culture for 30 days and analyzed again by flow cytometry.
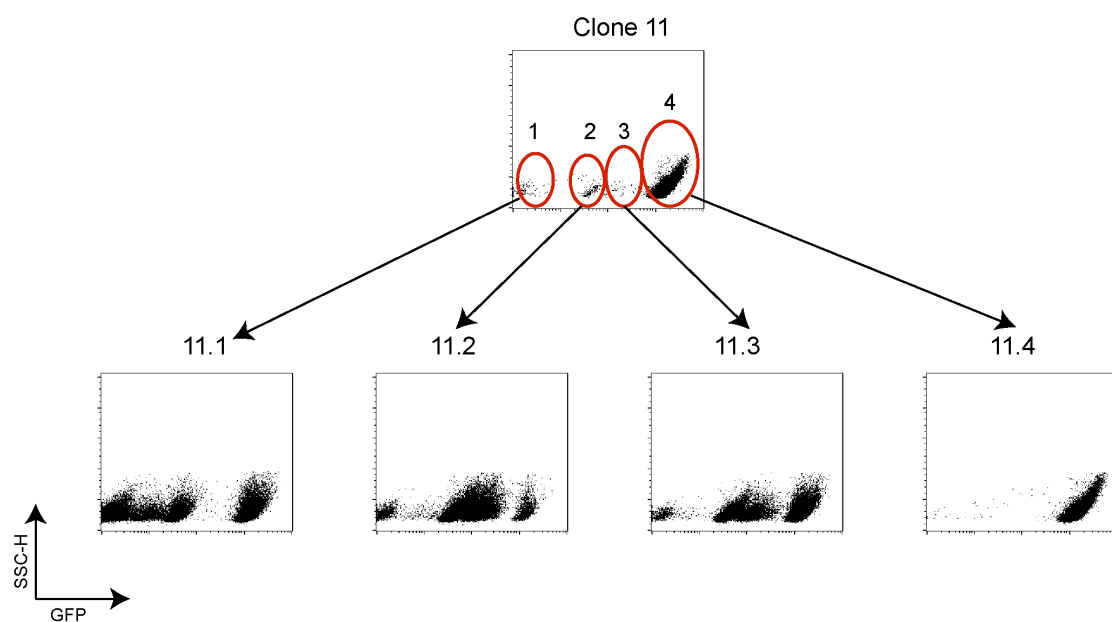
Fig. 47. Clone 11 sorted for four subpopulation, (upper blot) and analyzed by flow cytometry 30 days after sort (lower row). X-axis: GFP fluorescence on a logarithmic scale; Y-axis: side scatter.

As shown in Fig. 47, the subpopulation 11.4 did not lose its GFP expression over time. But the subpopulations 11.1, 11.2 and 11.3 behaved differently: regardless of the origin of the sort, they more or less reproduced the flow profile of the mother clone 11, i.e., they now contained cells of even higher GFP intensities than the ones we had sorted for. This could be due simply to a non-precise sort.

We considered three possibilities for low GFP expression: (i) silenced transcription, (ii) silenced translation, (iii) mutation in GFP. To distinguish between those, we isolated the intermediate population of 11.2 (designated as GFP intermediate) and the high population of 11.4 (designated as GFP high) by cell sorting, and immediately extracted DNA, RNA and Protein. To address the first possibility of silenced transcription, we ran a conventional RT-RCR and quantitative PCR on GFP transcripts (Fig. 48). As a result, we found a 2.44 fold higher steady-state level of GFP mRNA expression in the GFP high population as compared to the GFP intermediate population.
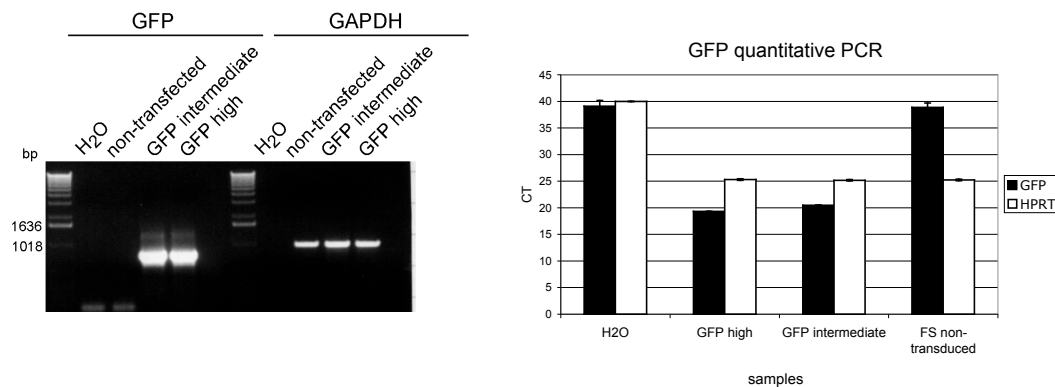
Fig. 48. Quantification of GFP expression: RT-PCR (left) and quantitative RCR (right) on RNA isolated of cells with intermediate GFP intensity (GFP intermediate) and cells with high GFP intensity (GFP high).

However, the intensity difference between GFP high and GFP intermediate is 10-fold (Fig. 47), and the 2.44 fold higher transcript in the GPF high (Fig. 48) may not completely account for that. Thus, we also tested for inhibition of translation by performing a Western with various amounts of protein lysate of the GFP high and GFP intermediate populations (Fig. 49).
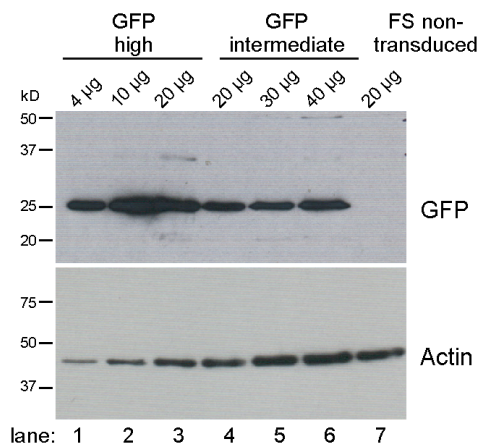
Fig. 49. GFP Western blot of whole cell lysates obtained from GFP intermediate and high populations. For quantification, different amounts of lysates were loaded. Actin as a loading control.

As a result, lane 1, loaded with 4 µg of GFP high lysate shows an intensity similar to the one of lane 6, loaded with 40 µg of lysate of GFP intermediate cells. From this we conclude that the difference in fluorescence intensity is reflected in the steady-state levels of GFP protein, and not by mRNA levels. There could well be an inhibition of translation.

So, it seemed that most of the changes intensity was not due to mutation. Nevertheless, to be sure that we did not miss any mutator action, we proceeded to sequence GFP of both populations from genomic DNA, and found that in both the stop codon TAG had reverted to TAC. In the intermediate population we found an additional C to T mutation resulting in an amino acid exchange from proline to serine. Because this is a drastic amino acid exchange (proline as a helix breaker), this mutation may contribute to decreased GFP intensity. While the mutations in the TAG stop codon are selected for by our method, the additional C to T mutation (originally) was not and thus may indicate an increased mutation rate. However, because of the complication of potential transcriptional silencing and translational inhibition we decided to shelve this line of experimentation. Nevertheless, before doing so, we determined the proviral integration site in clone 11.

Viral integration of clone 11

Using anchored PCR we determined the viral integration site of the FS clone 11 and found close to it the gene encoding Tid-1. Because Tid-1 had previously been identified as a potential interaction partner of AID in a yeast-two-hybrid screen (R. Harper, Wabl lab, unpublished data), we attributed some significance to the insertion side. Tid-1 is a member of the DnaJ family that contains a characteristic and conserved J domain. DnaJ proteins are cofactors for Hsp70 chaperones, and they are involved in DNA repair, which would also fit with a role in hypermutation. But if Tid-1 actually were involved in that process in the clone 11, the viral integration

would have to change the expression level of the Tid-1 gene. We therefore determined Tid-1 expression by conventional RT-PCR and by Western blot with samples from clone 11, GFP intermediate population (11.2) and clone 11 GFP, high population (11.4). In the mouse, three Tid-1 splice variants are known: a long, a short and an intermediate form. Fig. 50A shows various TID splice variants and the primer design: One primer combination amplifies Tid-1 long (Tid-1_L) and short form (Tid-1_S), the other detects intermediate Tid-1 (TID-1_I) only.
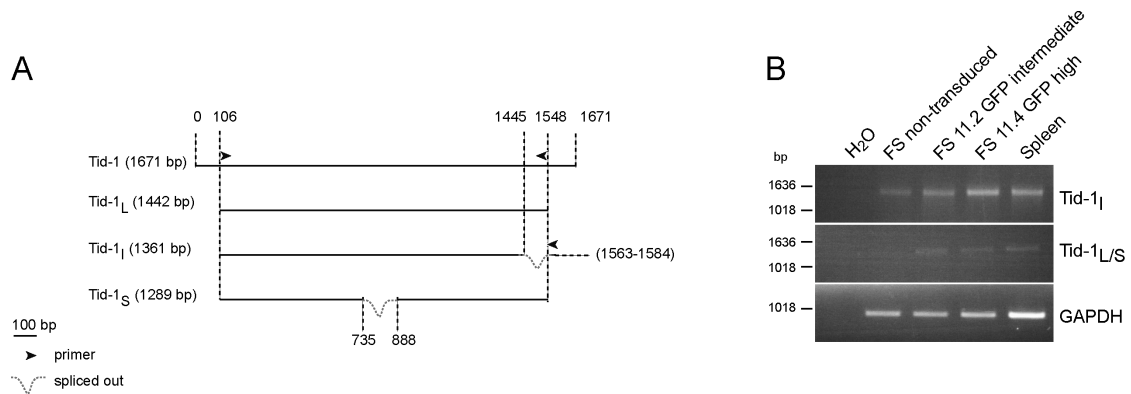
Fig. 50. (A) Primer design and three splice variants of Tid-1. Long (Tid-1_L), intermediate (Tid-1_I) and short Tid-1 (Tid-1_S) and their sizes indicated on the left. Numbers mark base pair positions. (B) Gelelectrophoresis with amplification products of the RT-PCR for different splice forms (right). RNA was extracted from the 11.2 intermediate population, and the 11.4 high GFP population; as well as from FS non-transduced and spleen cells as controls.

From the RT-PCR (Fig. 50), one may draw the conclusion that there is a slightly elevated transcription level of the Tid-$1_{L/S}$ mRNA in clone 11 descendants 11.2 and 11.4. Although both Tid-1 forms should have been amplified, we only detected the Tid-1_L form. Also, there may be increased transcript for the Tid-1_I splice form in the clone 11 subpopulations in comparison the FS non-transduced.
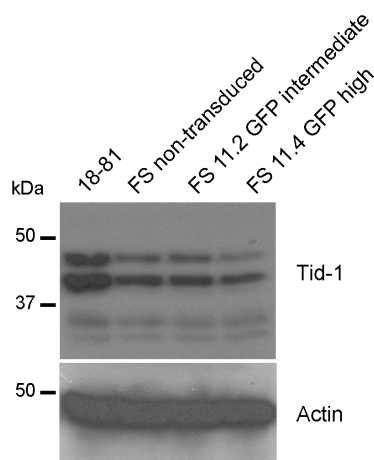
Fig. 51. Tid-1 Western blot with samples extracted from 11.2 intermediate population, 11.4 high GFP population, FS non-transduced as negative control, 18-81 as positive control. Actin as a loading control.

RT-PCR is notoriously unreliable as a method to quantify mRNA levels. Thus we assayed the steady-state levels of the Tid-1 protein. Fig. 51 shows a Western blot anti-human TID-1$_L$ as a first antibody. The bands at 40 and 43 kDa, respectively, most likely represent the two splice variants Tid-1$_L$ and Tid-1$_S$. Although the bands have the right sizes and the blot showed little background, we have no proof that the detected protein is Tid-1. Unfortunately, no negative control, for example in the form of a Tid-1 knock out cell, was available, and Tid-1 is a ubiquitously expressed protein. At any rate, if we take these two bands at face value to represent Tid-1, then there was no increase of Tid-1 protein in the clone 11 subpopulation over the FS non-transduced cells. Also, Tid-1 seems to be expressed at higher levels in 18-81.

Supt6h - a potential AID interaction partner

In the yeast two-hybrid screen using AID cDNA as bait, we have found three binding partners. Among them, Supt6h was isolated 11 times—3 times from the 18-81 library and 8 times from a human activated B-cell library (Fig. 52). Therefore, we further investigated Supt6h as a potential AID interaction partner.
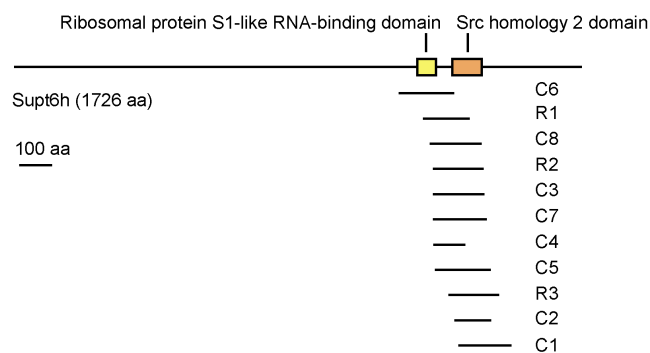
Fig. 52. Primary structure of Supt6h; lines underneath, 11 cDNA sequences recovered in the yeast two-hybrid screen. While the 5′ start sites of the sequences are precise, the 3′ ends of the cDNAs are undefined.

Endogenous expression of Supt6h in B cell lines

To address the question if and how much Supt6h is expressed in B cells, a Western blot was done with polyclonal rabbit anti-human Supt6h antiserum, which was raised against the very C-terminus of human Supt6h (Winkler et al. J. of Virology 2000) In Figs. 53A and B, the band at 200 kD represents Supt6h, which is expressed in the B cell lines 18-81, FS, and in NIH3T3.
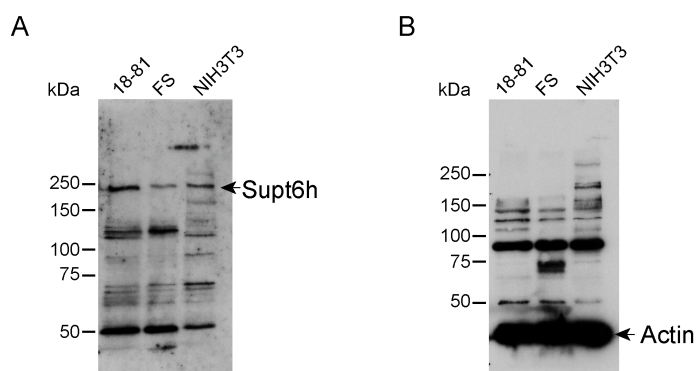
Fig. 53. Endogenous Supt6h expression. In all cells lines NIH3T3 is expressed. (A) Blot incubated with anti-Supt6h polyclonal serum. (B) Blot incubated with anti-actin antibody as a loading control.

Co-immunoprecipitation (Co-IP) of tagged AID and Supt6h

To confirm the interaction of AID and Supt6h, co-immunoprecipitations were performed. For these studies, tagged Supt6h and AID proteins were overexpressed in the human embryonic kidney cell line HEK 293. A schematic representation of the constructs is shown in Fig. 54. Tags were cloned into the N-terminus of the proteins.

Two tags, HA and FLAG, were added together to AID; either HA or Myc was added to Supt6h.
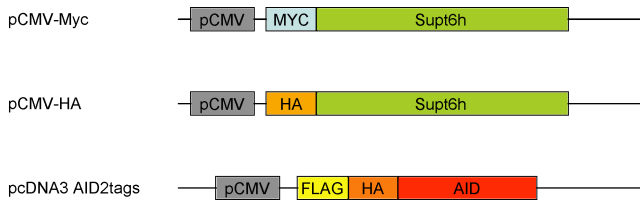

Fig. 54. Schematic representation of constructs encoding tagged Supt6h and AID, respectively. Supt6h has either a Myc or an HA tag at the N-terminus, whereas AID has both N-terminal FLAG and HA tag.

Vectors pCMV-Myc-Supt6h and pCMV-HA-Supt6h (Fig. 54) were transduced into HEK cells using the Lipofectamine method. Two days after transduction, whole cell lysate was prepared and a Western blot was run, shown in Fig. 56. The blots were incubated with different antibodies: (A) anti-Supt6h serum; (B) anti-Myc antibody; or (C) anti-Actin as a loading control.
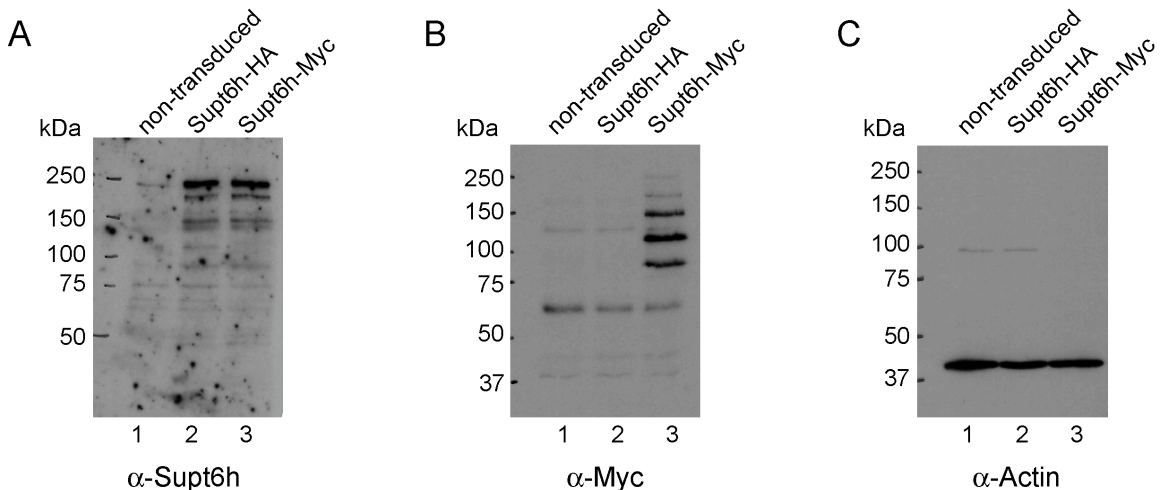
Fig. 55. Overexpression of Myc and HA tagged Supt6h in HEK cells. Cells transduced with the constructs denoted above the lanes were lysed, electrophoresed and Western blotted with the antibody indicated below the blots. Numbers to the left of the blots are molecular mass markers (in kDa).

As evident in Fig. 55, Myc and HA tagged Supt6h was successfully overexpressed in HEK cells. On blot A, lane 1, 'non-transduced', the band at 200 kD represents the endogenous Supt6h. In the lanes 2 and 3, respectively, both Myc- and HA-tagged constructs are well expressed. From blot B it seems that Supt6h is processed, either in the HEK cells directly or in their lysates. Because the Myc tag and the epitope to which the antibody was raised are on opposite ends, it looks as if the protein is cleaved or trimmed from the N-terminal end. In line with this are the other bands that stain with the antibody to Supt6h, which are smaller. But the protein seems to be

trimmed from the C-terminal end also, since there are lower-molecular-weight bands that strongly react with the antibody to the Myc tag.

The multiple bands might be interpreted as cross-reactivity, but in our case, they are clearly a product of the transduced (overexpressed) gene, which recapitulates the (very faint) endogenous expression pattern.

Immunoprecipition with AID-Flag

Since we wanted to get an indication of whether the exogenous AID and Supt6h products can bind to each other in the HEK cells, we precipitated, with antibody to FLAG, the lysates of HEK cells transduced with Supt6h-Myc, AID-FLAG, or both. The composition of the precipitates was then analyzed with antibodies raised against AID, the FLAG tag, Supt6h, or the Myc tag (Fig. 56). In lysates of the non-transduced HEK cells, antibody to Supt6h revealed endogenous Supt6h; anti-Myc tag stained endogenous Myc, at 65 kD, in addition to another (non-specific) band between 100 and 150 kD; and anti-FLAG stained a non-specific band above 75 kD. As expected, the anti-AID and anti-FLAG antibodies detected the exogenous AID in both lysates and precipitates, and the anti-Supt6h and anti-Myc antibodies detected exogenous Supt6h in the lysates. But these antibodies also stained exogenous Supt6h in the anti-FLAG precipitates; even though the anti-Supt6h staining is barely visible, the anti-Myc staining is quite strong.
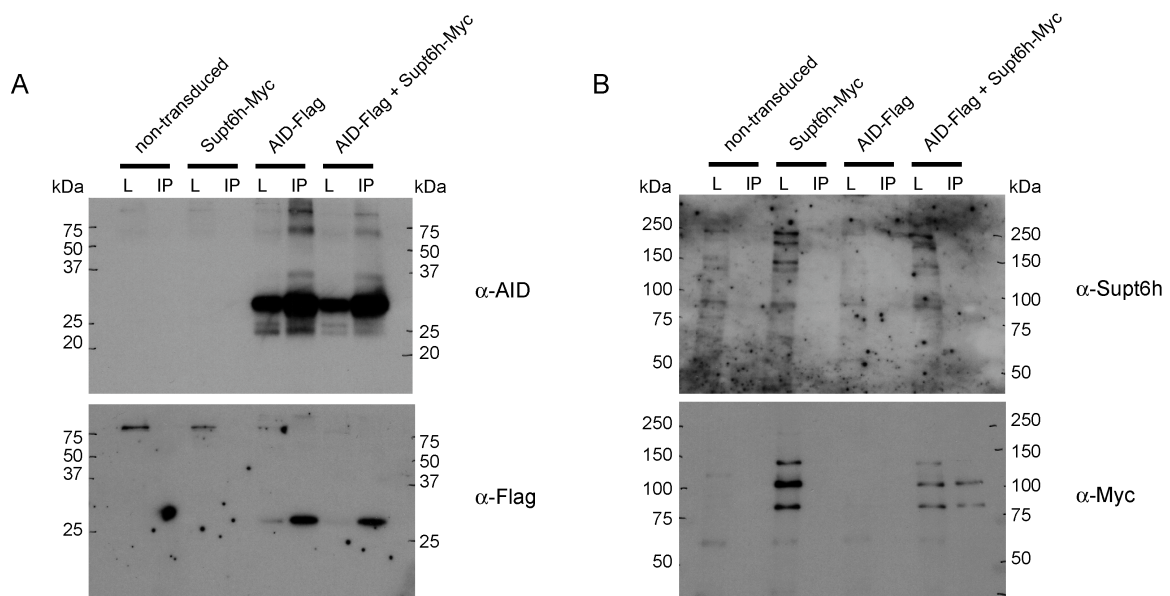

Fig. 56. Western blots of anti-FLAG-co-precipitated HEK cell lysates. Cells transduced with the constructs denoted above the lanes were lysed, electrophoresed and Western blotted with antibody indicated to the right of the blots. Numbers to the left and right of the blots are molecular mass markers (in kDa). On top of blots: transduced constructs; L = Lysate; IP = Co-Immunoprecipitation.

Immunoprecipition with Supt6h-Myc

Because of full-length Supt6h was not co-precipitated with AID-Flag, we tried to co-precipitate AID-Flag with Supt6h-Myc. Indeed, Fig. 57 shows that a 24-kD protein is pulled down in cells transduced with AID-Flag. However, this happened regardless whether Supt6h-Myc was co-transfected or not, suggesting AID binds to endogenous Myc, which is also precipitated with the anti-Myc coated beads.
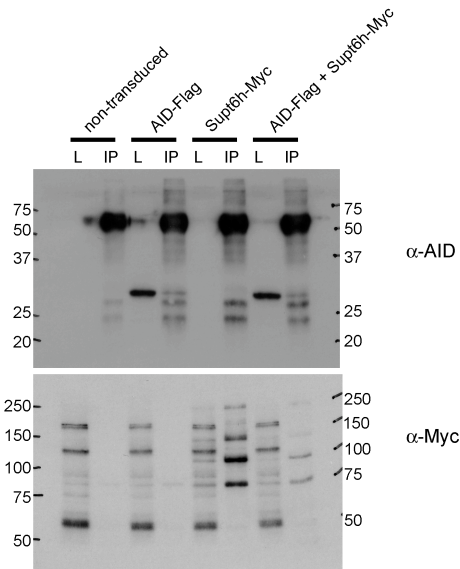

Fig. 57. Western blots of anti-Myc-co-precipitated HEK cell lysates. Cells transduced with the constructs denoted above the lanes were lysed, electrophoresed and Western blotted with antibody indicated to the right of the blots. Numbers to the left and right of the blots are molecular mass markers (in kD). L = Lysate; IP = immunoprecipitation.

Supt6h short hairpin RNA (shRNA)

Another way to prove the interaction of Supt6h with AID and therefore a function of Supt6h in CSR and/or SHM, is to silence its expression. Since Supt6h is a ubiquitous chromatin regulator and is known to be involved in embryonic development, we opted to do a conditional knock out. For this we used a lentiviral RNA interference vector (pSico) that allows for conditional activation via the Cre-lox system (Ventura et al., 2004). This vector, schematically represented in Fig. 58, is based on a strategy, in which the mouse U6 promoter has been modified by including a hybrid between LoxP site and TATA. Expression of Cre will lead to a recombination of the TATA-lox sites and result in the expression of a short hairpin RNA (shRNA), which then gives rise to RNAi. The recombination of the loxP sites is monitored by loss of YFP expression, and the viral vector encoding Cre has GFP as a marker.
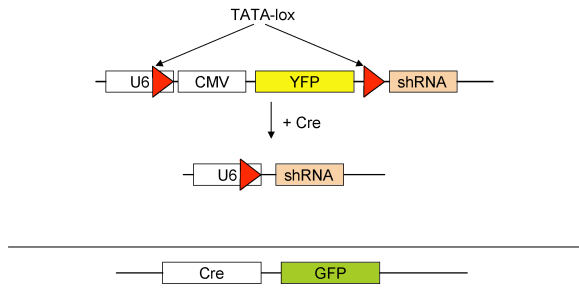
Fig. 58. Schematic presentation of the conditional shRNA construct with YFP as a marker (above). Upon recombination of the TATA lox sites mediated by Cre, shRNA is expressed and YFP expression lost. The vector encoding Cre (below) has GFP as a marker.

Three different shRNA, denoted #1 - #3, were designed to act on the 3' UTR of Supt6h and cloned in the pSico vector, and the cell line 18-81 was transduced with each of the three separately. Because 5 days after transduction only 5-8% of the cells were YFP positive, we sorted for YFP positives. Fig. 59 shows a flow cytometry analysis 4 days after the sort. As can be seen in the right plot, 93% of the transduced cells contained the conditional siRNA construct.
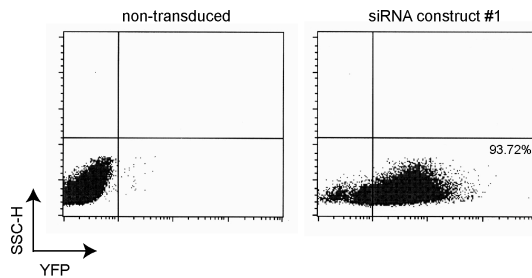

Fig. 59. Flow cytometry analysis 18-81 transduced with the conditional shRNA lentivirus containing shRNA #1. Y-axis: side scatter; X-axis: GFP.

For expression of the hairpins, cells were transduced with a second virus (MIG), encoding the Cre recombinase, and GFP as a marker. The initial transduction efficiency was 34% two days after transduction (Fig. 60)
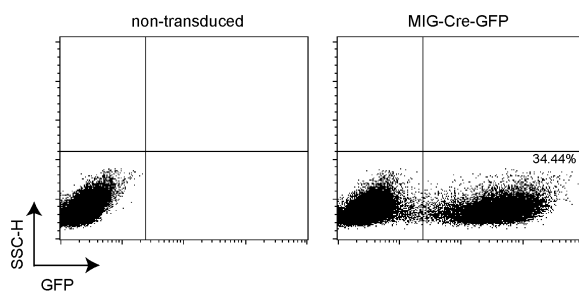

Fig. 60. Flow cytometry analysis of 18-81 transduced with the MIG encoding Cre recombinase, and GFP as a marker. Y-axis: side scatter; X-axis: GFP.

29 days after transduction with the shRNA-encoding lentivirus, cells were transduced with the Cre-GFP encoding MIG virus. Three days after transduction with MIG, cells were sorted for expression of shRNA, which are GFP positive and YFP negative cells (see above). Fig. 61 shows the dot plot representation of the FACS. As expected, cells upon expression of Cre lose YFP expression.

non-transduced   Cre only

shRNA#1   shRNA#1 + Cre

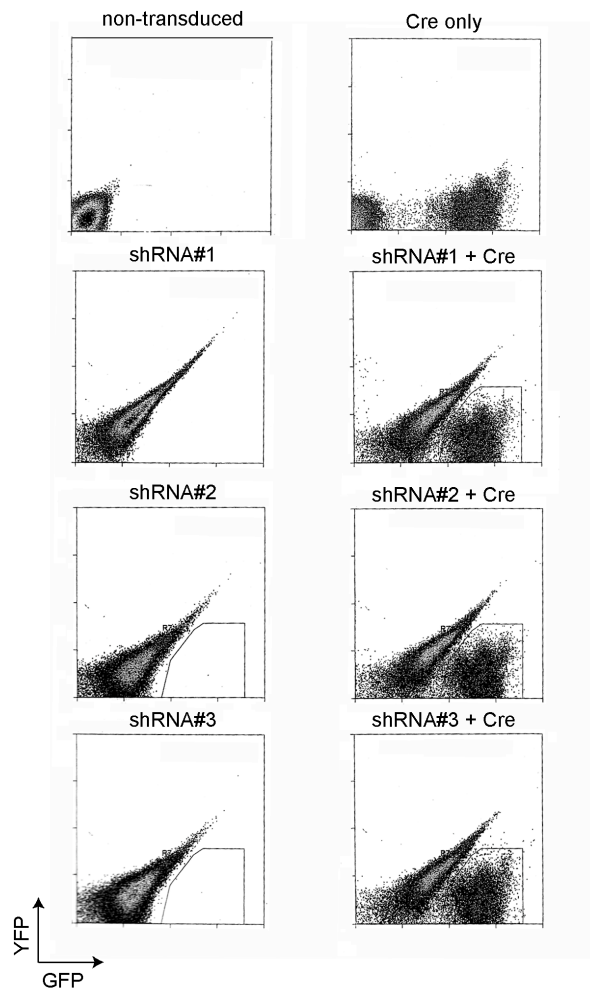shRNA#2   shRNA#2 + Cre

shRNA#3   shRNA#3 + Cre

YFP

GFP

Fig. 61. FACS of 18-81 transduced with 3 different conditional shRNA constructs only (left panel); or with an additional virus encoding Cre (right panel). Transducing constructs are denoted on top of plots. Y-axis: YFP; X-axis: GFP. Because YFP and GFP emission spectra are close together, high compensation had to be used.

To check the efficiency of the silencing of Supt6h via shRNA, protein from the various transduced cells shown in Fig. 61 was extracted and run in a Western blot with anti-Supt6h antiserum (Fig. 62). The band at 200 kD represents Supt6h. Unfortunately, although the transductions and marker expressions were as supposed, there was no significantly decreased Supt6h expression, for any of the shRNA constructs. Perhaps the shRNA construct #3 has a slight effect on Supt6h expression. But for sure this reduction would not be sufficient for the evaluation of Supt6h function in hypermutation and class-switching.

RNAi construct: non-transduced   #1   #2   #3
Cre: -   -   +   -   +   -   +

kDa
250   α–Supt6h
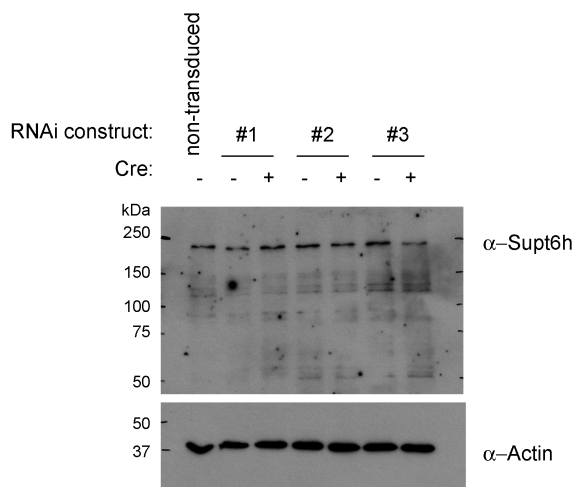150
100
75

50

50   α–Actin
37

Fig. 62. Western blot from 18-81 transduced with various siRNA constructs, and with or without Cre (indicated above blot). Actin as a loading control. Numbers to the left of the blot are molecular mass markers (in kDa). First antibody/serum use is indicated to the right of the blot.

It is known that most of such strain-specific differences occur in the UTR. So far, the only sequence published for murine Supt6h was obtained from the mouse strain Swiss Webster, and this is the sequence we based our design on for the shRNAs we have used. Since SNPs in the target sequence for shRNAs can have drastic effect, the 3'UTR of BALB/c, which is the genetic background of 18-81, has to be sequenced. Based on this sequence, new shRNAs will have to be designed. Therefore, we are now designing slightly different constructs, and we also take into account differences in sequences between mouse strains.