Research article

Cloning and functional characterisation of avian transcription factor E2A

Thomas M Conlon and Kerstin B Meyer*

Address: Cambridge Institute for Medical Research and Department of Pathology, University of Cambridge, WT/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2XY, UK

Email: Thomas M Conlon - tmc33@cam.ac.uk; Kerstin B Meyer* - kbm1000@mole.bio.cam.ac.uk

* Corresponding author

Abstract

Background: During B lymphocyte development the E2A gene is a critical regulator of cell proliferation and differentiation. With regards to the immunoglobulin genes the E2A proteins contribute to the regulation of gene rearrangement, expression and class switch recombination. We are now using the chicken cell line DT40 as a model system to further analyse the function of E2A.

Results: Here we report the cloning and functional analysis of the transcription factor E2A from chicken. Using RACE PCR on the chicken lymphoma cell line DT40 we have isolated full-length clones for the two E2A splice variants E12 and E47. Sequence conservation between the human and chicken proteins is extensive: the basic-helix-loop-helix DNA binding domain of human and chicken E47 and E12 are 93% and 92% identical, respectively. In addition high levels of conservation are seen in activation domain I, the potential NLS and the ubiquitin ligase interaction domain. E2A is expressed in a variety of tissues in chicken, with higher levels of expression in organs rich in immune cells. We demonstrate that chicken E12 and E47 proteins are strong transcriptional activators whose function depends on the presence of activation domain I. As in mammals, the dominant negative proteins Id1 and Id3 can inhibit the function of chicken E47.

Conclusions: The potential for homologous recombination in DT40 allows the genetic dissection of biochemical pathways in somatic cells. With the cloning of avian E2A and the recent description of an *in vitro* somatic hypermutation assay in this cell line, it should now be possible to dissect the potential role of E2A in the regulation of somatic hypermutation and gene conversion.

Background

The transcription factor E2A contributes to transcriptional regulation in many cell lineages. However, it is essential for the development of B lymphocytes [1,2]. Its role in mammalian B cell development has been studied extensively and E2A functions in B cell commitment and proliferation as well as immunoglobulin (Ig) gene rearrangement and expression (reviewed in [3]). By alternative splicing the E2A gene encodes two basic helix-loop-helix (bHLH) proteins, E12 and E47, which differ only in their highly homologous DNA binding and dimerisation domains [4]. Binding sites (consensus CANNTG) for these transcription factors are found in all Ig enhancers as well as a number of genes required for heavy and light chain rearrangement (*λ5, rag-1, rag-2, EBF*). Remarkably, over-expression of E2A together with the recombinase activating genes is sufficient to allow rearrangement of the endogenous Ig locus in a non-lymphoid cell line [5].

E2A also plays a role in peripheral B cell differentiation. E2A protein expression is greatest in the highly proliferative dark zones of germinal centres, where class switching (CSR) and somatic hypermutation (SHM) are thought to occur. In keeping with such a role, repression of E2A via Id proteins inhibits CSR [6] and leads to lower expression of the AID (activation induced deaminase) gene [7], a gene that is essential for both CSR and SHM to occur [8,9].

Avian B cell development differs from mammalian B cell development in a number of important aspects. Rearrangement occurs between a single V_L and J_L segment to yield a single functional variable light chain gene. Similarly a single V_H segment combines with one of 15 D and a single J segment [10] generating only a small repertoire. Gene conversion subsequently utilises a pool of upstream pseudo V-genes to generate diversity [10,11]. Thus gene conversion is the primary mechanism to establish the B cell repertoire in chickens. The diversification and expansion of B cell progenitors occurs in the specialised microenvironment of the Bursa of Fabricus from where mature B cells exit into the periphery. By 6–8 months the diversity of the B cell compartment is established and the bursa involutes. Given the distinct nature of avian B cell development we were interested to find out whether the transcription factor E2A plays a similar central role in B cell ontogeny and the generation of diversity.

To this end we have now cloned full-length avian E12 and E47 cDNAs from the chicken B cell lymphoma DT40 and have initiated their functional characterisation.

Results and Discussion
Cloning of chicken E12 and E47
A chicken bursal EST database [12,13], was searched with the bHLH domain of human E47, revealing a single highly homologous clone of 785 bp. Using a combination of degenerate and RACE PCR a 2.47 kb sequence encoding full-length chicken E47 was obtained. The protein sequence downstream of the first methionine is highly conserved across human, mouse, xenopus and chicken (Fig. 1A) suggesting that the true 5' end has been identified. Furthermore, nucleotides 3 bp and 6 bp upstream of the ATG correspond to the Kozak consensus sequence [14]. In order to address whether chicken, like its mammalian homologues, contains the alternatively spliced E12 DNA binding exon, primers were designed on either side of the bHLH region (P7, 8) and used in RT-PCR on DT40 cDNA. The resultant clones contained either E12 (542-620aa) or E47 (542-616aa) sequence (boxed sequences in Fig. 1A and 1B), while the sequence flanking the bHLH exon was identical.

Sequence analysis
The resultant full-length E12 and E47 protein sequences were compared to the human, mouse and xenopus orthologues using CLUSTAL W [15], and the results of this analysis are depicted in Fig. 1. Chicken E12 and E47 are identical except for their bHLH domain, which is nevertheless highly homologous with 72% identity and 86% similarity when allowing only conservative changes. Compared to their human orthologues, chicken E47 and E12 are 71% and 69% identical, respectively. Within the bHLH domains the conservation is even greater. Chicken and human E47 bHLH are 93% identical and chicken and human E12 bHLH are 92% identical. Thus E47 bHLH domains of different species are more similar than the E12 and E47 domain within one species. With regard to the other species examined, chicken E12 is slightly more homologous to xenopus E12 (70%) than to mouse E12 (68%). Outside the bHLH domain there are a number of well-conserved domains. In particular, activation domain I (ADI) is highly homologous across all four species examined (see Fig. 1). The chicken sequence corresponds to the helix consensus described for interaction with the SAGA complex [16], indicating that E2A mediates some of its function at the chromatin level. Activation domain II has been identified as a domain required to drive transcription in insulin producing β-cells [17]. Within activation domain II only the 5' region is conserved, while insertions are found within the 3' region. However the length of the conserved region, which is significantly longer than that of the ADI, may indicate that this domain has additional functions. The putative nuclear localisation domain is identical in all the compared sequences. Another functional region of the E47 protein is the domain interacting with ubiquitin conjugating enzyme UbcE2A (477-530aa [18,19]). Within this region, runs of very high homology are seen between the sequences. Whether this is sufficient to obtain functional interaction with UbcE2A awaits experimental confirmation.

In conclusion all regions of very high sequence conservation correspond to functionally important domains.

Genomic structure of the E2A gene
We have mapped the intron-exon structure of the E2A gene on the most recent release of the chicken genome using the genome browser at UC Santa Cruz [20]. Chicken E2A is encoded on chromosome 28 and comprises 19 exons spanning 38.7 kb. A schematic representation of the gene is shown in Fig. 2. The intron-exon boundaries have been mapped (Table 2, see Additional file 1) and all introns conform to the splice consensus sequence. Three exons (2, 5 and 7) lie in gap sequences of the genome and searches with these exons did not reveal any homologies in the chicken genome. Their exact location has therefore not been assigned. However, the length of these three

A Alignment
   of E12

```
human    -MNQP-QRMAPVGTDKELSDLLDFSMMFPLPVTNGKGRPASLAGAQFGG-SGLEDRPSSG  57
mouse    MMNQS-QRMAPVGSDKELSDLLDFSMMFPLPVANGKSRPASLGGTQFAG-SGLEDRPSSG  58
xenopus  -MNQ-QQRMTAVGTDKELSDLLDFSMMFPLPVANGKTRPTTLASTQFGGGSGLEERTGSA  58
chicken  -MNQQQQRMAAVGTDKELSDLLDFSMMFPLPVANGKNRPTTLASTQFGG-SGLDERPGSG  58
         ***  ***:.**:.******************.*** **::*..**.* ***::*..*.
              QR    K L D LDFS    P   K
                     helix of ADI

human    SWGSGDQSSSSFDPSRTFSEGTHFTES-----HSSLSSSTFLGPGLGGKSGERGAYASFG  112
mouse    SWGSSDQNSSSFDPSRTYSEGAHFSDS-----HSSLPPSTFLGAGLGGKGSERNAYATFG  113
xenopus  PWGTEDHNGSTFDQGRGYGDGTHYGEHRELPSHEGNLSAPFLASGLVGKN-DRPPYSTFG  117
chicken  SWGTGDQNSSTFDQGRSYGEGPHYGEHRDLPSHNSLSSSPFLGPGLVGKSGERASYSTFG  118
         .**: *:..*:** .* ::.*.*: :     *.. .:.**..** **. :* .*::**

human    RDAGVGGLTQAGFLSGELALNSPGPLSPSGMKGTSQYYPSYSGSSRRRAADGSLDTQPKK  172
mouse    RDTSVGTLSQAGFLPGELSLSSPGPLSPSGIKSSSQYYPSFPSNPRRRAADGGLDTQPKK  173
xenopus  RDSGMT-MNQTGFLSGDLAMNSPSALSPNAGKAGSQYY-TYPNNSRRRPGDSNLDAQPKK  175
chicken  RDAGMPGLSQTGFLPSEMGIASPSTLSPTGVKGGSQYF-SYPNNPRRRGAESSIDGQPKK  177
         **:.:  :.*:***.::.:  **..***.. *. *** :::...*** .:::* ****

human    VRKVPPGLPSSVYPPSSGEDYGRDATAYPSAKTPSSTYPAPFYVADGSLHPSAELWSPP-  231
mouse    VRKVPPGLPSSVYPPSSGDSYSRDAAAYPSAKTPSSAYPSPFYVADGSLHPSAELWSTP-  232
xenopus  VRKVPPGLPSSVYPANSGDEYCGDRGSYASTKTPNSVYPGAFYMADG-LHSS-DIWASSG  233
chicken  VRKVPPGLPSSVYPSNSGDDYSRDPAGYTPSKPPSTVYPGAFY**M**ADG-LHNSPDLWSSPG  236
         ***************.**:.*  *  .*:.:*.*.:**.:***  ** ** ::*:..
         NLS

human    --GQAGFGPMLGGGSSPLPLPPGSGPVGSSGSSSTFGGLHQHERMGYQLHGAEVNGGLPS  289
mouse    --SQVGFGPMLGDGSSPLPLAPGSSSVGSG----TFGGLQQQDRMGYQLHGSEVNGSLPA  286
xenopus  GIGQGGYSSVLNSSQSSVSQTTG------------FSSLHPHERMGFPMHSGEVN---PS  278
chicken  AMSQSSYGAMLGSSSSPLPQSAS------------FNSLHQHMNYQLHSGELNGGLPS  284
          .* .:.:*....*.:  ...            *..*: ::**.: :*..*.:*   *:

human    ASS-FSSAPGATYGGVSSHTPPVSG--ADSLLGSRGTTAGSSGDALGKALASIYSPDHSS  346
mouse    VSS-FSAAPG-TYSGTSGHTPPVSGAAAAESLLGTRGTTASSSGDALGKALASIYSPDHSS  344
xenopus  VTSSFSSTPA--QYGVSSHTPPIST--GDTIIGNRGTATGSSGDALGKALASIYSPDHSS  334
chicken  VSG-FSSAST--PYGVSSHTPPISG--TDTMMGNRGTTAGSSGDALGKALASIYSPDHSS  339
         .:. **:.   *.*.****:*   :::*.***:*.:**************.*********
                                                          ADII

human    NNFSSSPSTPVGSPQGLAGTSQWPRAGAPGALSPSYDGGLHGLQSKIEDHLDEAIHVLRS  406
mouse    NNFSPSPSTPVGSPQGLPGTSQWPRAGAPSALSPNYDAGLHGLQSKMEDRLDEAIHVLRS  404
xenopus  TNFSSTPSTPVGSPQGITGSGQWPRANGPGALSPSYDGTLHTLQNKMEDRLDEAIHVLRS  394
chicken  NNFSSNPSTPVGSPQGLAGTSQWSRAGGQGALSPSYEGSLHTLQNKMEDRLDEAIHVLRN  399
         .***.*********** *:*..**.**..  .****.*:. ** **.*:**:*********.

human    HAVGTAG-------DMHTLL-------PGHGALASGFTG-PMSLGGRHAGLVGGSHPEDG  451
mouse    HAVGTAS-------DLHGLL-------PGHGALTTSFTG-PMSLGGRHAGLVGGSHPEEG  449
xenopus  HAVGQTGALPGGHDDIHSLL------SATASVLGSGFPAAVLSLANRHSMLN--SHTEEA  446
chicken  HAVGQTAAMPSNHGDMHGLLGSAPAHSAAVGSLGQAFPASVMSLGGRHPSLVGGGHPEDG  459
         **** :.        *:* **       .  *  .*.  .:**.**.* .*.*.:

human    LAGSTSLMHNHAALPSQPGTLPDLS-RPPDSYS---G-LGRAGATAAASEIKREEKEDEE  506
mouse    LTSGASLLHNHASLPSQPSSLPDLSQRPPDSYS---G-LGRAGTTAGASEIKREEKEDEE  505
xenopus  LPSASNMLHSQVTVPAQPGSLPDLT-RPQDSYSGLPG-LGRAIPPGRVPDIKRESKEDEE  504
chicken  LSSTPSMLHNHVTLPSQPSSLPDLN-RQQDTYSGLSGGLGRSSVSSGTNEIKREEKEDEE  518
         *.. ..::*.:.::*:**:**.****. * *:**    * ***: .. . :***.*****
                                              Ubc-L

human    NTSAADHSEEEKKELKAPRART-----SPDEDE---DDLLPPEQKAEREKERRVANNARE  558
mouse    IASVAD-AEEDKKDLKVPRTRT-----SPDEDE---DDLLPPEQKAEREKERRVANNARE  556
xenopus  NRSVADLSDDEKKESKPQRSRTRCSINSQDEDEDEDDDNLPPEQKAEREKERRVANNARE  564
chicken  NTSVADNSEEEKKELKPSRNRT-----SQDEDEE--DDLLPPEQKAEREREKRRVANNARE  571
         *.** :::.:**: *  * **     *  ****   ** ********* .*********

human    RLRVRDINEAFKELGRMCQLHLNSEKPQTKLLILHQAVSVILNLEQQVRERNLNPKAACL  618
mouse    RLRVRDINEAFKELGRMCQLHLSSEKPQTKLLILHQAVAVILSLEQQVRERNLNPKAACL  616
xenopus  RLRVKDINEAFKELGRMCQLHLNSEKPQTKLLILHQAVSVILSLEQQVRERNLNPKAACL  624
chicken  RLRVRDINEAFKELGRMCQLHLNSEKPQTKLLILHQAVSVILNLEQQVRERNLNPKAACL  631
         ****:***************** :***********************:***.*******************
                              E12 bHLH exon

human    KRREEEKVSGVVGDPQMVLSAPHPGLSEAHNPAGHM  654
mouse    KRREEEKVSGVVGDPQLALSSAAHPGLGEAHNPAGHL  652
xenopus  KRREEEKVSGV--DPQMGLSGGHPGVGDSHNPVGHM  658
chicken  KRREEEKVSGVVGDPQMTLSASHPGLGDGHNPVGHM  667
         **********  ***: **. ***::.:***.**:
```

B Alignment
   of E47
   bHLH exon

```
human    530 STDEVLS-------LEEKDLRDRERRMANNARERVRVRDINEAFRELGRMCQMHLKSDKA  582
mouse    528 STDEVLS-------LEEKDLRDRERRMANNARERVRVRDINEAFRELGRMCQLHLKSDKA  580
xenopus  533 -QDEDEDEDSKEDCVEEKDTKDRERRMSNNARERVRVRDINEAFKELGRMVQMHMKADKA  591
chicken  542 STDEALS-------LEDKDLRDRERRMANNARERVRVRDINEAFKELGRMCQMHLKTDKA  594
             **  .     .:*:** .:******.*************.****** *:*:*:***

human    QTKLLILQQAVQVILGLEQQVR  604
mouse    QTKLLILQQAVQVILGLEQQVR  602
xenopus  QTKLIILQQAVAVIMSLEQQVR  613
chicken  QTKLIILQQAVQVILGLEQQVR  616
         ****:****** **:.******
```

Figure 1
Alignment of E2A sequences from human, mouse, xenopus and chicken: full-length E12 (**A**) and the E47 bHLH exon (**B**) which replaces the E12 bHLH exon in the full-length sequence. "*" denotes identical residues, ":" conserved substitutions as recorded in CLUSTAL, and "." semi-conservative substitutions. Accession numbers for the sequences used are: PI5923, XM125750, X66959 for human, mouse and Xenopus, respectively. The chicken sequence has been assigned the accession number AJ579995 for E12 and AJ579996 for E47. Shaded boxes denote functionally defined regions: activation domain I (ADI) with boxed helix [29], activation domain II (ADII) [17] and the ubiquitin ligase interaction domain [18]. The putative NLS is underlined and the bHLH region is boxed in black. The exon border of the xenopus sequence is based on its homology to the other sequences. Chicken amino acid 221 (bold) is the first aa of the ΔADI constructs used in the functional analysis.
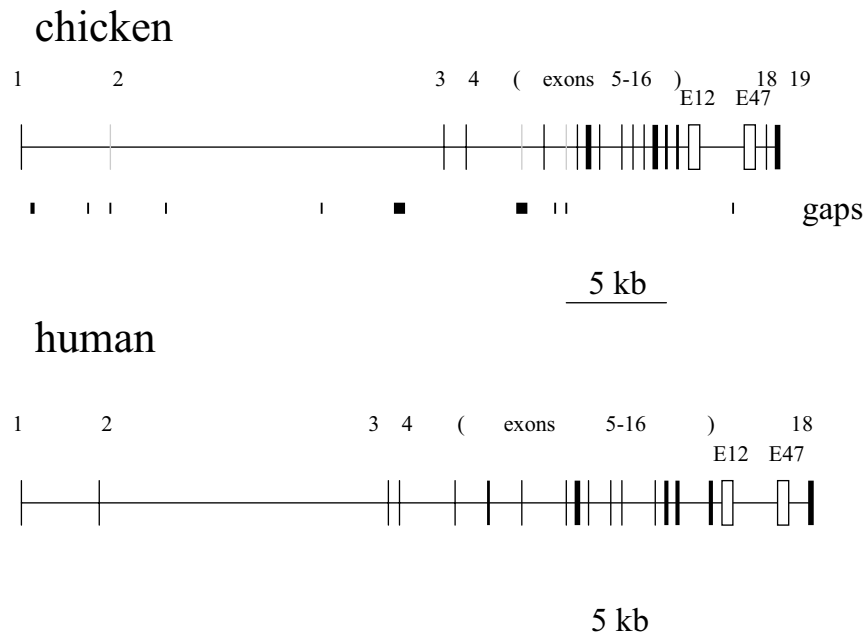
Figure 2
Schematic representation of the structure of the chicken and human E2A genes. Black lines represent assigned exons, grey lines represent exons lying in gap sequences; their exact position has not been determined. Open boxes represent the alternatively spliced E12 and E47 exons encoding the helix-loop-helix domain of these proteins. See Table 2 for the exact position of each of the exons.

exons is very similar to exons 2, 5 and 7 in the human gene (see Table 2) suggesting that they are indeed encoded as single exons. The overall structure of the human and chicken genes is extremely well conserved, with the human gene containing 18 exons spanning 40.98 kb of chromosome 19 (Fig. 2). The intron-exon borders are very similar, the only exception being the 3'UTR which is split into 2 exons in the chicken whilst the 3' UTR is encoded as a single exon in the human gene. In both chicken and human the first 2 exons are widely spaced, with exons 3 lying 21.5 kb upstream of the ATG in the chicken and 17.9 kb in the human E2A gene. In contrast the exons at the 3' end of the gene are highly clustered with exons 8–19 spanning 9.2 kb in the chicken and 10.4 kb in the human gene.

Furthermore we found three clusters of highly related sequence between the chicken and human loci that did not correspond to coding sequence. These mapped to upstream and downstream of the gene and to a region upstream of exon 3. We are currently investigating whether these function as regulatory sequences.

Chicken E2A expression pattern
To examine whether there is differential expression between E12 and E47 in the chicken DT40 cell line, splice variant specific primers (P10, P11 and P12) were used in semi-quantitative RT-PCR of DT40 mRNA. Fig. 3A demonstrates that the relative amounts of E12 and E47 expressed are very similar. Furthermore we examined E2A gene
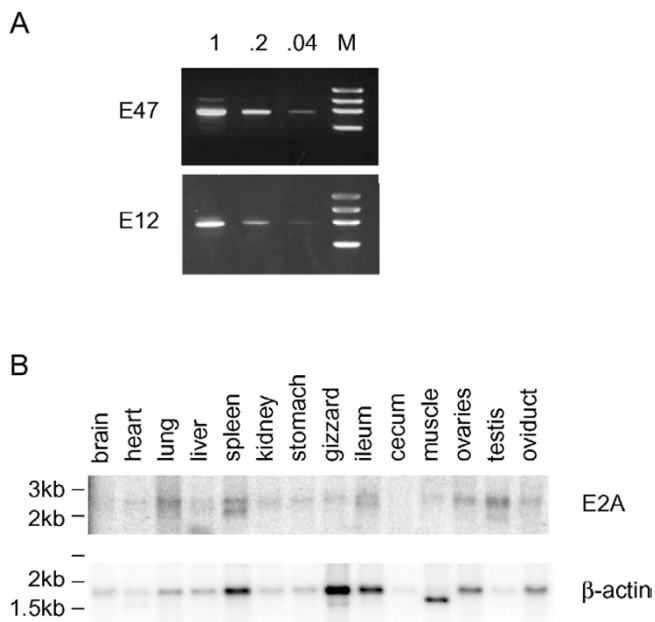
A

B

Figure 3
A. PCR analysis of E12 and E47 on DT40 cDNA using increasing dilutions of cDNA. M, DNA markers. **B.** Northern Blot of E2A expression, with 2 μg polyA+ RNA from the tissues shown above the lanes. A β-actin probe was used as a control.

expression in different chicken parenchymal tissues. As observed in human and mouse, low levels of E2A expression are detected by Northern blotting in many different tissues in the chicken (Fig. 3B). Of the tissues examined, greatest expression is detected in the lung, spleen and ileum, sites where immune cells accumulate. However, relatively high expression is also seen in testis and the ovaries. The size of the E2A message is approximately 2.5 kb which corresponds well to the length of cDNA clones we have isolated. However, in spleen and to a lesser extent in ileum and testis there is a second smaller message of about 2.2 kb. The functional significance of this is not yet clear. The overall expression pattern is in keeping with E2A contributing to gene expression in a variety of tissues but being of particular importance for lymphocytes. Like in mammalian cells the wide expression pattern may not reflect protein levels since E2A is regulated post-transcriptionally [6,21].

Functional analysis of chicken E2A
Having cloned chicken E2A we wished to confirm its ability to activate transcription from E2A containing reporter plasmids. We demonstrate that chicken E47 and E12 strongly activate transcription from a multimerised E2A

binding site linked to a TATA box (Fig. 4A). Upon transfection into human HEK293 cells, hamster E47 (MDE47 [22]) stimulated transcription by 1000-fold. Chicken E12 and E47 gave similar values with E12 being approximately 20% less active than E47. For both isoforms transcriptional activation strongly depended on the presence of its 5' end, including activation domain I (ADI). E12ΔADI retains less than 1% of the activity of the full-length protein, while E47ΔADI is reduced to 3.5% of its wild type activity. The deleted region encompasses the ADI and the putative NLS. However, previous studies have shown that similar deletions of human E2A are nevertheless found in the nucleus [17]. This suggests that the loss of function in the mutant proteins is due to the deletion of activation domain I. Furthermore we investigated whether Id proteins, dominant negative regulators of bHLH function, can regulate chicken E2A proteins. To this end we amplified mouse Id1 [23] and mouse Id3 [24] by PCR and subcloned the resultant fragments into an expression vector, driven by the cytomegalo virus promoter/enhancer region, and tested their ability to abrogate transcriptional activation (Fig. 4B). Once again E2A transactivation was assayed on a multimerised E2A binding site. Under the experimental conditions used this yielded 4500-fold activation by E47. As before E12 activity was about 80% of that seen for E47. Co-transfection of mouse Id1 reduces E47 driven transcription to 10% of that seen for the control vector only, while the mouse Id3 was able to reduce transcription to 26%. Thus, like their mammalian orthologues chicken E2A proteins can be regulated by the Id proteins. Orthologues of the Id proteins in chicken have been cloned and their embryonal expression pattern examined [25]. As yet little is known about their expression in avian lymphoid development. However, our results suggest that Id proteins modulate the activity of chicken E2A proteins.

Conclusions
Here we report the cloning of the chicken orthologue of the human transcription factor E2A. The identification of chicken orthologues of genes known to be important for the development of the immune system is of great interest with regard to understanding a related but distinct immune system as is found in birds. One of the advantages of utilising chicken as an experimental system is the existence of the bursal cell line DT40, which undergoes homologous recombination with high frequency thus allowing a genetic approach to the dissection of biochemical pathways. We plan to utilise DT40 in the analysis of E2A function in mature B cells. In mammalian cells, E2A is known to contribute to the regulation of class switch recombination and it has been suggested that it may contribute to SHM either directly or indirectly through controlling AID gene expression. Normally IgV gene diversification in chicken occurs by gene conversion,
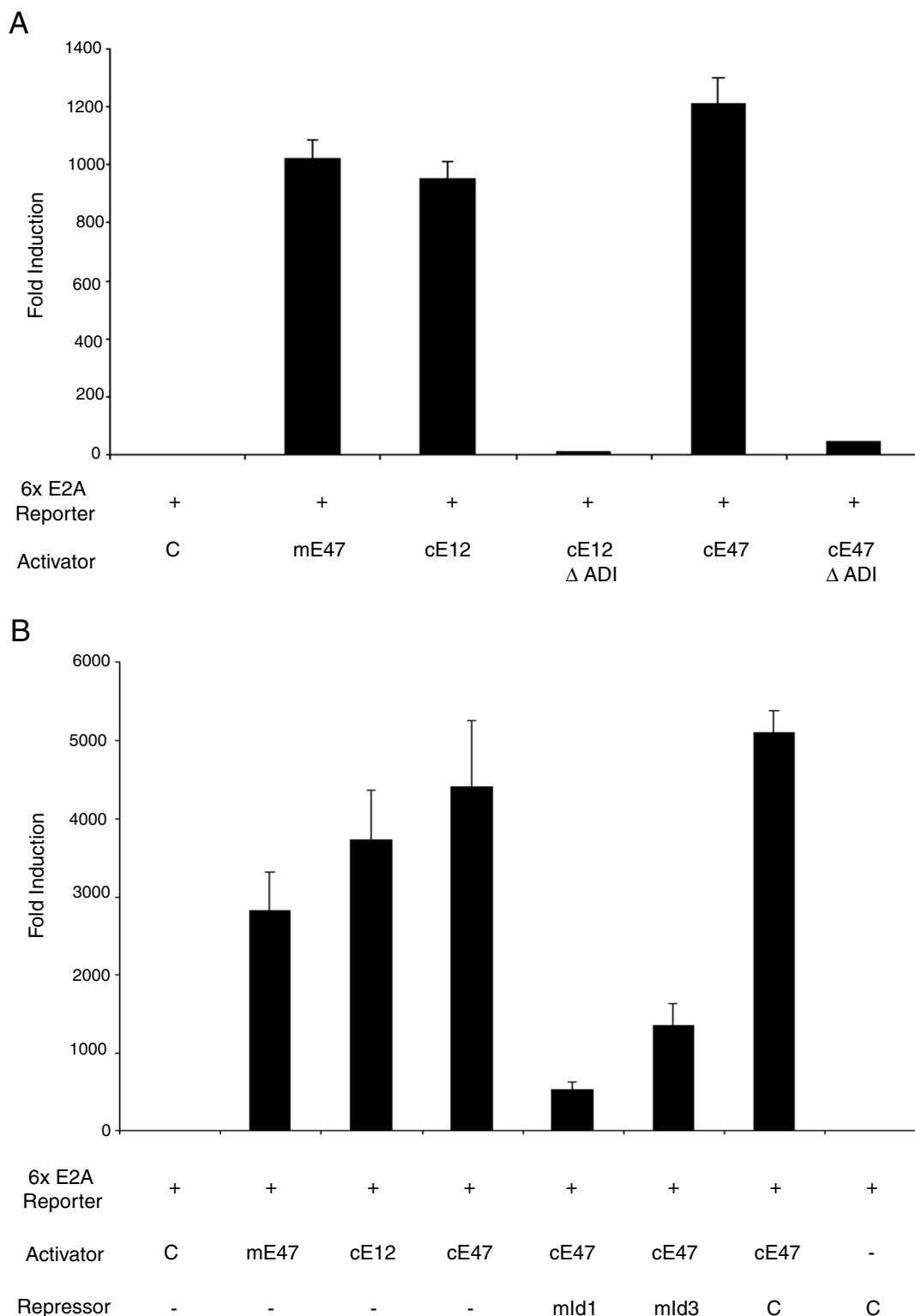
Figure 4
Transcriptional activation by mammalian E47 and chicken E12 and E47 in HEK293 cells. **A.** Transactivation of a multmerised E2A binding site by different E2A proteins as indicated. Values are given relative to those obtained with the pcDNA control vector. The data shows the mean +/- SEM of triplicate transfections. **B.** Effect of co-transfection of Id expression constructs on E47 transcriptional activation. The data show the mean +/- SEM of five independent transfections.

Table 1:

Primer name	Sequence, 5' to 3'
P1	CTGGCCCACGGCGTGATTCCTCAAG
P2	GAAAGTTCCGGAGTGGTGGGAGACC
P3	GGGGTTTACTGCGAAAAGTCGAAG
P4	AGTGACCTCCTGGACTTCAG
P5	GGATCCCTGCTGTAGTCATCACCTGAG
P6	CTCCAAACTGAGTGCTTGCTAGGGTCG
P7	CAGGGATCCGGCTGGATAC
P8	AAGCTTAGGATCTTTACATGTGTCC
P9	GGTACCCAGGACTGCCACTGTTGAG
P10	CTACCAGCTGCATTCAGGAGAG
P11	TCCTCCTCATCCTCATCTTG
P12	GATAAGGCCTCGTCTGTACTG
P13	ATGAAGGTCGCCAGTGGCAG
P14	TCAGCGACACAAGATGCG
P15	GTGCTGTGTTCCCATCTATCGT
P16	TGGACAATGGAGGGTCCGGATT
P17	ATGAAAGTCGCCAGTGGCAG

Sequence of primers used in PCR reactions.

although somatic mutation can occur [26]. However, cells carrying mutations in the Rad51 homologues XRCC2 and XRCC3 undergo SHM with very high frequency [27]. The identification of avian E2A, together with the recent description of an in vitro somatic mutation assay in DT40 cells, should allow us to investigate the exact role E2A plays in regulating the complex events that confer specificity on the SHM mechanism.

Methods
RACE PCR and sequencing
Primers P1, P2 and P3 (for all primers see Table 1) were used in 5' and 3' RACE PCR as described in the SMART RACE cDNA Amplification Kit (BD Biosciences, USA) on DT40 total RNA (RNeasy, Qiagen, UK). Further PCR reaction with primers P4, P5 and P6 generated full-length E47 cDNA. All PCR products were cloned into the pCR2.1-TOPO vector (Invitrogen, USA) and sequenced using the ABI Big Dye terminator cycle sequencing ready reaction kit (Applied Biosystems, UK). Sequencing reactions were analysed in an ABI 3700 DNA Analyser (Applied Biosystems, UK).

E2A expression
PCR reactions on serial dilutions of DT40 cDNA prepared using the SMART RACE cDNA Amplification Kit, were carried out using primer pairs P10 and P11 for E12 expression and P10 and P12 for E47 expression. 2 μl of cDNA was used in 20 μl reaction mixtures containing 0.2 mM of each dNTP, 2 μl 10x PCR buffer (Applied Biosystems, UK), 1.5 mM $MgCl_2$ (E12) or 2 mM $MgCl_2$(E47), 1 μM of each primer and 0.5 U Taq Gold (Applied Biosystems,

UK). PCR conditions were 94°C for 10 min followed by 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 3 min, followed by a 10 min extension at 72°C. Products were analysed on a 1.5% agarose gel. A NeverFail Northern Blot (RNway Laboratories) of 2 μg chicken poly A+ RNA per lane was probed with a ^{32}P-labelled fragment of E2A genomic DNA, overlapping the two exons from 1454–1576 and 1577–1726 of the cDNA. The β-actin control probe was generated by PCR using primers P15 and P16.

Plasmids
Full length cDNAs for expression cloning of E12 and E47 were generated by a 2-step PCR reaction using primer pairs P7, P8 (3' end of proteins) and P9, P5 (extending to 5' start). Clones were confirmed by sequencing and cloned into the pcDNA3 expression vector (Invitrogen, USA) using *Bam*H I and *Not* I to generate ΔAD1 E12 and E47 constructs. Both E12 and E47 are truncated at amino acic 221 (indicated in bold in figure 1). Subsequently the 5' ends were inserted using *Kpn* I and *Bam*H I to generate full length clones. The MDE47 vector was obtained from M. Sigvardsson and contains a forced dimer of hamster E47 in pcDNA [22]. The 6xE2A luciferase vector (pXp2Luc) from A. Green's laboratory contains 3 copies of the following sequence: gtcgaa<u>cagat</u>gttcacacgac<u>catct-gtgg</u>. pGL3-control and promoter vectors are from Promega, UK. Mouse Id1 was amplified using the primer pairs P13, P14 and cloned into the pCR2.1-TOPO vector. Id1 was subcloned into pcDNA3.1/Hygro expression vector (Invitrogen, USA) using *Bam*H I and *Not* I. Full length mouse Id3 [28] was subcloned into the same expression vector using *Bam*H I and *Hind*III.

Cell culture and transfection
The chicken bursal B cell line DT40 cells was obtained from Dr J. Sale and maintained in RPMI-1640, 50 μM β-mercaptoethanol, 7% foetal calf serum, 3% chicken serum and antibiotics; human embryonic kidney cells, HEK293, were from Dr K.J. Patel and maintained in DMEM, 10% foetal calf serum and antibiotics. 1 μg each of luciferase reporter and transactivation plasmids, and 0.25 μg of CMV-β-galactosidase vector, were transfected using Superfect Reagent (Qiagen, UK) according to the manufacturer's instructions. We harvested cells 36–48 hours after transfection and generated cell extracts in 200 μl Promega reporter lysis buffer. Luciferase and β-galactosidase activity in 5 μg of protein was measured using Promega reagents. Results are given as ratios of luciferase over β-galactosidase activity.

Authors' contributions
TMC performed the cloning, sequencing and expression studies. KBM initiated the project, carried out some of the functional and the sequence analysis and drafted the

manuscript. All authors have read and approved the final manuscript.

Additional material

Additional File 1

Table 2 Location of the chicken E2A exons, and exon and intron sizes of the chicken and the human E2A genes.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2172-5-11-S1.doc]

Acknowledgements

References

1. Bain G, Maandag ECR, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC, Krop I, Schlissel MS, Feeney AJ, van Roon M, van der Valk M, te Rile JPJ, Berns A, Murre C: **E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements.** *Cell* 1994, **79:**885-892.
2. Zhuang Y, Soriano P, Weintraub H: **The helix-loop-helix gene E2A is required for B cell formation.** *Cell* 1994, **79:**875-884.
3. Kee BL, Quong MW, Murre C: **E2A proteins: essential regulators at multiple stages of B-cell development.** *Immunol Rev* 2000, **175:**138-149.
4. Murre CP, McCaw PS, Baltimore D: **A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, myoD, and myc proteins.** *Cell* 1989, **56:**777-783.
5. Romanov WJ, Langerak AW, Goebel P, Wolvers-Tettero ILM, van Dongen JJM, Feeney AJ, Murre C: **E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells.** *Mol Cell* 2000, **5:**343-353.
6. Quong MW, Harris DP, Swain SL, Murre C: **E2A activity is induced dring B-cell activation to promote immunoglobulin class switch recombination.** *EMBO J* 1999, **18:**6307-6318.
7. Sayegh CE, Quong MW, Agata Y, Murre C: **E-proteins directly regulate expression of activation-induced deaminase in mature B cells.** *Nat Immunol* 2003, **4:**586-593.
8. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T: **Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme.** *Cell* 2000, **102:**553-563.
9. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A: **Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2).** *Cell* 2000, **102:**565-575.
10. Reynaud CA, Dahan A, Anquez V, Weill JC: **Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region.** *Cell* 1989, **59:**171-183.
11. Reynaud CA, Anquez V, Grimal H, Weill JC: **A hyperconversion mechanism generates the chicken light chain preimmune repertoire.** *Cell* 1987, **48:**369-378.
12. Abdrakhmanov I, Lodygin D, Geroth P, Arakawa H, Law A, Plachy J, Korn B, Buerstedde JM: **A large database of chicken bursal ESTs as a resource for the analysis of vertebrate gene function.** *Genome Res* 2000, **10:**2062-2069.
13. [http://swallow.gsf.de/dt40est.html]: **.** .
14. Kozak M: **The scanning model for translation: an update.** *J Cell Biol* 1989, **108:**229-241.
15. [http://www.ebi.ac.uk/clustalw]: **.** .
16. Massari ME, Grant P, Pray-Grant MG, Berger SL, Workman JL, Murre C: **A conserved motif present in a class of helix-loop-helix**
17. proteins activates transcription by direct recruitment of the SAGA complex.** *Mol Cell* 1999, **4:**63-73.
17. Aronheim A, Shiran R, Rosen A, Walker MD: **The E2A gene product contains two separable and functionally distinct transcription activation domains.** *Proc Natl Acad Sci USA* 1993, **90:**8063-8067.
18. Kho CJ, Huggins GS, Endege WO, Hsieh CM, Lee ME, Haber E: **Degradation of E2A proteins through a ubiquitin-conjugating enzyme, UbcE2A.** *J Biol Chem* 1997, **272:**3845-3851.
19. Loveys DA, Streiff MB, Schaefer TS, Kato GJ: **The mUBC9 murine ubiquitin conjugating enzyme interacts with the E2A transcription factor.** *Gene* 1997, **201:**169-177.
20. [http://genome.ucsc.edu/index.html]: **.** .
21. Meyer KB, Mufti DAH: **Post-transcriptional regulation of E2A proteins via lipopolysaccharide and CD40 signaling.** *Eur J Immunol* 2000, **30:**719-724.
22. Sigvardsson M, O'Riordan M, Grosschedl R: **EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes.** *Immunity* 1997, **7:**25-36.
23. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H: **The protein Id: a negative regulator of helix-loop-helix DNA binding proteins.** *Cell* 1990, **61:**49-59.
24. Christy BA, Sanders LK, Lau LF, Copeland NG, Jenkins NA, Nathans D: **An Id-related helix-loop-helix protein encoded by a growth factor-inducible gene.** *Proc Natl Acad Sci USA* 1991, **88:**1815-1819.
25. Kee Y, Bronner-Fraser M: **Id4 expression and its relationship to other Id genes during avian embryonic development.** *Mech Dev* 2001, **109:**341-345.
26. Parvari R, Ziv E, Lantner F, Heller D, Schechter I: **Somatic diversification of chicken immunoglobulin light chains by point mutations.** *Proc Natl Acad Sci U S A* 1990, **87:**3072-3076.
27. Sale JE, Calandrini DM, Takada M, Takeda S, Neuberger MS: **Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation.** *Nature* 2001, **412:**921-926.
28. Meyer KB, Skogberg M, Margenfeld C, Ireland J, Pettersson S: **Repression of the immunoglobulin heavy chain 3' enhancer by helix-loop-helix protein Id3 via a functionally important E47/E12 biding site: implications for developmental control of enhancer function.** *Eur J Immunol* 1995, **25:**1770-1777.
29. Quong MW, Massari ME, Zwart R, C. Murre.: **A new transcriptional activation motif restricted to a class of helix-loop-helix proteins is functionally conserved in both yeast and mammalian cells.** *Mol Cell Biol* 1993, **13:**792-800.