Context dependent function of APPb enhancer identified using enhancer trap-containing BACs as transgenes in zebrafish

Leighcraft A. Shakes[1], Tennison L. Malcolm[1,2], Kevin L. Allen[1,3], Supriyo De[4], Ken R. Harewood[1,3] and Pradeep K. Chatterjee[1,2,*]

[1]Julius L. Chambers Biomedical/Biotechnology Research Institute, [2]Department of Chemistry, [3]Department of Biology, North Carolina Central University, Durham, NC 27707 and [4]Gene Expression and Genomics Unit, National Institute on Aging, NIH, Triad Technology Center, Baltimore, MD 21224, USA

ABSTRACT

An enhancer within intron 1 of the amyloid precursor protein gene (APPb) of zebrafish is identified functionally using a novel approach. Bacterial artificial chromosomes (BACs) were retrofitted with enhancer traps, and expressed as transgenes in zebrafish. Expression from both transient assays and stable lines were used for analysis. Although the enhancer was active in specific nonneural cells of the notochord when placed with APPb gene promoter proximal elements its function was restricted to, and absolutely required for, specific expression in neurons when juxtaposed with additional far-upstream promoter elements of the gene. We demonstrate that expression of green fluorescent protein fluorescence resembling the tissue distribution of APPb mRNA requires both the intron 1 enhancer and ~28 kb of DNA upstream of the gene. The results indicate that tissue-specificity of an isolated enhancer may be quite different from that in the context of its own gene. Using this enhancer and upstream sequence, polymorphic variants of APPb can now more closely recapitulate the endogenous pattern and regulation of APPb expression in animal models for Alzheimer's disease. The methodology should help functionally map multiple noncontiguous regulatory elements in BACs with or without gene-coding sequences.

INTRODUCTION

About two-thirds of the highly conserved genome sequence between human and other vertebrates as divergent as the fish does not code for proteins, is distributed throughout the genome, and mostly located at large distances along the DNA from the start sites of genes (1–7). Part of these conserved noncoding elements (CNEs) plays a role in regulating gene expression, and is believed to be essential to all vertebrate development (2,4–6). Conservation of gene regulatory function has also been demonstrated recently in the absence of sequence similarity (8), suggesting that structural features of DNA can be preserved despite their different sequence. Despite these findings, tools for functionally analyzing CNEs continue to use a 'targeted approach', where PCR amplified CNE–DNA is joined to a reporter gene and analyzed for expression either in mice or zebrafish (9,4,10). Thus amplified CNE–DNA, was either coinjected with linear reporter DNA (5,6) or introduced as reporter vector plasmids (4) into zebrafish eggs, and analyzed for transient expression of green fluorescent protein (GFP) fluorescence. A second approach used the Tol2 transposon system that allowed CNE-reporter gene fusions to be integrated into the germline more efficiently (10). While such studies have greatly enhanced our understanding of CNE function, and can be scaled up, they encounter hurdles when multiple regulatory domains from noncontiguous DNA act in concert to regulate expression of the gene. Difficulties also arise when the noncoding regulatory DNA is not conserved across species and thus not recognizable prior to testing. A third approach used the traditional enhancer trap with a pseudo-typed murine leukemia virus to infect dechorionated zebrafish embryos (11) to identify regulatory sequences in a nontargeted fashion, but more importantly, in the context of the gene and chromosome. However, subtractive analysis requiring deletion of sequences thought to act in combinatorial fashion with other noncontiguous enhancing elements, remains a hurdle with this approach. The task of functionally identifying regulatory DNA of either the conserved or nonconserved

variety therefore presents a real challenge, and is likely to benefit most from using a 'nontargeted' approach that is also unbiased. Because such regulation is often observed over large distances along the DNA, using bacterial artificial chromosomes (BACs) and P1-derived artificial chromosomes (PACs) (12–14) might prove most beneficial, as issues of context of those regulatory modules to the gene can be addressed simultaneously.

A novel approach that addresses several of these issues has been developed. An enhancer trap comprising of a basal promoter driven reporter gene, such as GFP, is positioned at short intervals along the genomic DNA in the BAC clone with the help of a Tn10 transposon. The set of enhancer trap modified BACs is then introduced into zebrafish eggs, and the patterns of GFP expression used to map *cis*-acting gene regulatory elements functionally in the BAC DNA.

We have explored the role of conserved as well as nonconserved DNA in regulating expression of the amyloid precursor protein (APP) gene, that is central to Alzheimer's disease (AD), because the understanding of its regulation remains incomplete (15–21). An enhancer with novel properties within intron 1 of the gene has been identified in this report using the new approach. We demonstrate that it is required, along with ∼28 kb of DNA upstream of the gene, for expression of GFP fluorescence resembling the tissue distribution of APPb mRNA in neuronal cells reported earlier (22). This newly identified enhancer has unique characteristics of tissue specificity: when operating with its full complement of far upstream regulatory elements, the expression is exclusively in neurons, but its enhancing function is restricted to the notochord when acting through its basal promoter elements. These observations have implication for APP gene expression and, more generally, for context-dependent function of *cis*-regulatory sequences (23,24).

MATERIALS AND METHODS

Four BAC clones from the zebrafish genomic library, CH211-235E22, CH211-219P8, CH211-192O20, CH211-43O16 designated here as BACs A, B, C and D, respectively, were purchased from BAC/PAC resources, Oakland, CA, USA. DNA fragments from the zebrafish genome are in the pTARBAC2.1 vector in these clones. End deletions of insert DNA in the BAC clones with loxP transposons were generated using procedures described earlier (25,26). Procedures for DNA isolation/purification from BAC deletions, field inversion gel electrophoresis (FIGE) analysis, end sequencing of BAC deletions with transposon-based primers have been described earlier (27,28). Identical procedures were followed with the BAC deletions generated with enhancer trap transposons. DNAs from BAC clones A, B, C were used as templates to amplify segments of DNA from intron 1 of APPb, and they gave identical products. Primer sequences used for amplifying APPb intron 1 enhancer were:

LF2: 5′ d CACCTGAATGTGGGATTTGGTTG 3′
LR4: 5′ d GATAGACTCTGGCAATCTG 3′

LF1: 5′ d CCAAAGCATTCTTCTGAG 3′
LR5: 5′ d CAAGAGTCTTGGCTCCATGTG 3′

Basal promoter fragments of 0.75 and 0.35 kb were amplified from zebrafish APPb BAC DNA using the following primers:

UF1: 5′ d AGCCATATCCTGTATATAG 3′
UR1: 5′ d CTGTGTTCCCAAGCGCAGCAC 3′
UF3: 5′ d AGAGGTAGGTTGAGGCAACAATAAC 3′

The PCR amplified DNA from LF2-LR4 or LF1-LR5 for the intron element (IE), or UF1-UR1 and UF3-UR1 for the basal promoter elements 0.75 and 0.35 kb, respectively, was cloned in pCR2.1 vector before being constructed into the enhancer trap transposons.

Zebrafish egg injections

Typically ∼20–50 pg of highly purified circular BAC DNA, prepared by the Qiagen Sciences, Maryland, USA column procedure described earlier (28), was injected into each egg using an injection station from World Precision Instruments, Florida, USA and a Nikon SMZ1500 microscope. The developing embryos were scored and analyzed for GFP fluorescence after 48 h, using a Nikon Diaphot Fluorescence microscope with a Nikon high pressure mercury lamp as the excitation source, and photographed with a RT Spot from Diagnostic Instruments, Michigan, USA.

RESULTS

Regulated expression of APP gene in appropriate tissue is important for building suitable animal models for AD because a 42-amino acid peptide expressed from the gene is found in neurofibrilatory tangles as amyloid plaques in brains from AD patients (29). However, to date expression of APP gene using its endogenous promoter elements is unavailable (29,30). It led us to conclude that key regulatory sequences remained unidentified, and BACs containing the gene were used so as not to preclude regulation from distal promoter elements.

A loxP transposon procedure to make progressive end deletions in BACs has been used to map genetic markers and gene regulatory elements both on a physical map of the chromosome (27,31), as well as functionally using cell lines or transgenic mice (32,33). An important feature of this technology is its ease of determining exactly where in the BAC the loxP-transposon had inserted to create the truncation (27,28). An additional feature is the ability to introduce reporter genes and other DNA cassettes precisely at the new end created in the large BAC clone: sequence in front of the loxP arrowhead as shown in Figure 1A is retained after the recombination event that creates the deletion (note orientation of arrow refers to directionality of loxP sequence). It is this particular feature that we have now utilized to place a basal promoter containing GFP gene in the BAC, such that potential regulatory elements further upstream (shown as RE-1, -2 and -3 in Figure 1A) can drive reporter gene expression when the retrofitted BAC DNA is introduced into zebrafish as a transgene. DNA sequences +0.147 to −0.75 kb
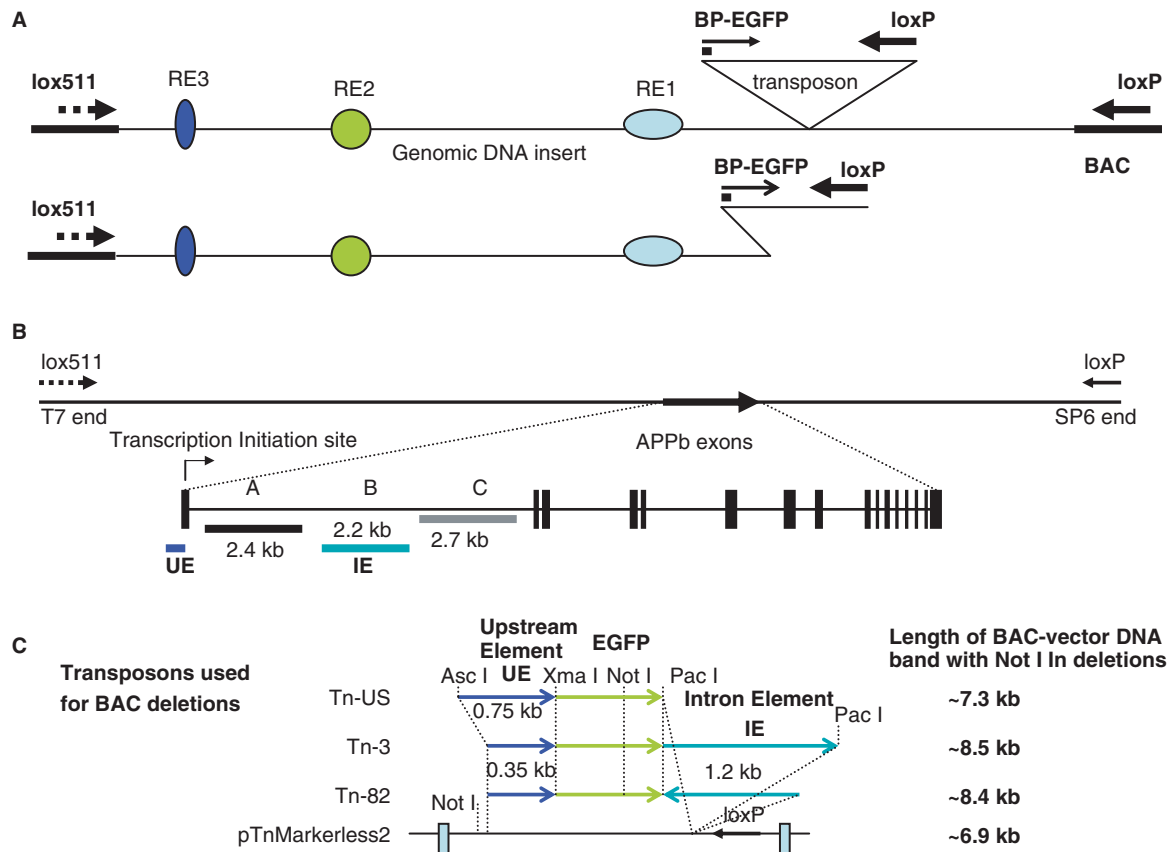
Figure 1. (**A**) Schematic representation of the methodology. A loxP-Tn10 is used to deliver a basal promoter GFP enhancer trap (indicated as BP-EGFP on top left end of inverted triangle) into the BAC DNA, which is then expressed as a transgene. Possible upstream regulatory elements (marked as RE-1, -2 and -3) are indicated by the colored circles/elipses. The genomic insert DNA in the BAC is flanked by loxP and lox511 sites indicated by the thick arrow and thick broken arrow, respectively. Cre-recombination of the loxP site endogenous to the BAC and the transposed loxP site deletes the genomic DNA from the right end, thus placing at that end the EGFP enhancer trap contained in the transposon. (**B**) Fine structure of APPb gene showing exons as thick vertical black bars. UE shown in blue comprises DNA segment +0.147 to −0.35 kb (i.e. 0.147 kb of UTR, +0.35 kb of sequence immediately upstream of APPb transcription start site, marked by the thin arrow at start of exon 1), and is part of all enhancer trap transposons except Tn-US, Tn-10 and Tn-13, where it is +0.147 to −0.75 kb. Sections of intron 1 DNA, 2–3 kb in length, A–C, were PCR amplified and tested for enhancer activity by constructing them into transposon plasmids. Fragment B, the 2.2 kb intron 1 region colored dark green and indicated as IE tested positive, while fragments A and C tested negative. Enhancer activity of fragment B was narrowed down to 1.2 kb in Tn-82. (**C**) Schematic drawing of Tn-US, Tn-3 and Tn-82. Note Tn-3 and Tn-82 contain the ~1 kb IE (dark green arrow) in opposite orientations. The blue arrow and the light green arrow represent UE and EGFP, respectively, with the arrowheads indicating direction of reading frame. The transposon plasmid pTnMarkerless2 has been described in (74). The UE (0.35 or 0.75 kb) is flanked by AscI and XmaI sites, and the IE (1.2–2.2 kb, or 2.4 and 2.7 kb in various constructs) is flanked by PacI sites. The size of BAC-vector DNA bands generated with Not I enzyme digests of DNA from BAC deletions serve as a diagnostic for authentic Cre-mediated loxP–loxP recombination, and is shown to the right of each of the transposons Tn-US, Tn-3 and Tn-82.

or +0.147 to −0.35 kb surrounding the transcription initiation site of APPb comprised the basal promoter in the enhancer trap transposons used in this study (Figure 1B and C).

Enhancers of transcription have traditionally been identified through transient expression of reporter genes in small plasmids (34–39). Transient expression from an episomal plasmid is suitable because the gene/enhancer(s) in stably integrated DNA in cell lines are often influenced by chromosomal regions adjoining it. This leads to variable expression in different lines. Thus although few of such episomal expression studies yield insights into transcriptional mechanisms that involve chromatin remodeling/modification, they appear capable of identifying enhancer elements rapidly. Transcription enhancement

by CNEs has also used transient expression in zebrafish (4–6), or mice (9,24), and similar assays are used here to identify enhancers in APPb. Stable transgenic zebrafish lines derived from our enhancer trap BACs also show variable expression in neurons (see below). Thus, enhancer assays used here have relied more on transient expression, although the major conclusions derived from these are supported by the data from germline transgenic fish.

Characterization of zebrafish APPb BACs

The BAC clones (160–200 kb size) used here were characterized by sequencing the ends of progressive deletions of insert DNA using transposon-based primers Seq 1 or

Seq 4 as described earlier (27). BLASTs to the zebrafish genome at Ensembl (Zv7) were consistent with there being no major rearrangements within inserts of BACs A–D, at the resolution of ∼1 kb of FIGE.

No expression from BACs deleted with Tn-US that carries only 0.75 kb upstream sequence (UE) fused to GFP as enhancer trap

The exon–intron structure of APPb gene is shown in Figure 1B. LoxP transposons were constructed with only the +0.147 to −0.75 kb DNA around the transcription initiation site of APPb [(shown as upstream element (UE)] fused to the GFP gene, shown as Tn-US in Figure 1C. A set of deletions in APPb BAC clone C, ranging in insert DNA size from 20 to 90 kb was made with this transposon. A FIGE analysis of the BAC-C deletions with Tn-US is shown in Figure 2.

Linear transposon DNA from Tn-US was first injected into zebrafish eggs, but no fluorescence was observed (data not shown). Next, the Qiagen purified DNA from a set of six BAC deletions made with Tn-US were injected. No fluorescence was observed (data not shown). Location of Tn-US enhancer trap in each BAC deletion was confirmed by end sequencing. The Tn-US generated enhancer trap BACs had up to 90 kb of DNA upstream, but no APPb-specific DNA downstream of +0.147 kb of the gene. Clearly a *cis*-acting sequence downstream, critical to expression of APPb, was missing in these BAC constructs.

Comparative genome sequence analysis in the APPb gene region

A cross species genome sequence pair-wise comparison of this region was conducted. Peaks of highly conserved DNA between zebrafish and human (Figure 3, second row), or to a lesser degree in Fugu (first row), but not as conspicuous in the mouse (third row), was found within intron 1 of APPb (marked by arrowhead within vertical broken green lines in Figure 3). The genome is expanded 4-fold for human and mouse but compressed 2-fold for Fugu in the region analyzed in Figure 3. Although sequence conservation within intron 1 is low in the mouse (third row between green lines), conservation of function might still exist (8) [see also (40,41)]. Therefore, the possibility of highly conserved (human and Fugu with zebrafish), and semi-conserved (mouse with zebrafish) regions within intron 1 enhancing expression of APPb seemed plausible, and was next tested.

Transposon DNA alone, with 3 cassettes [intron 1 (IE) + GFP + 0.75 kb upstream sequence, UE], elicits a specific pattern of GFP expression in the notochord of zebrafish when injected

Being unable to express APPb with only the ∼0.75 kb UE, the role of semi-conserved DNA within intron 1 was explored by incorporating these sequences in the enhancer trap. Because transposition efficiency of Tn10 drops off severely with very large inserts (42), a systematic narrowing down of the minimum essential segment capable of enhancer activity within the ∼10 kb intron 1
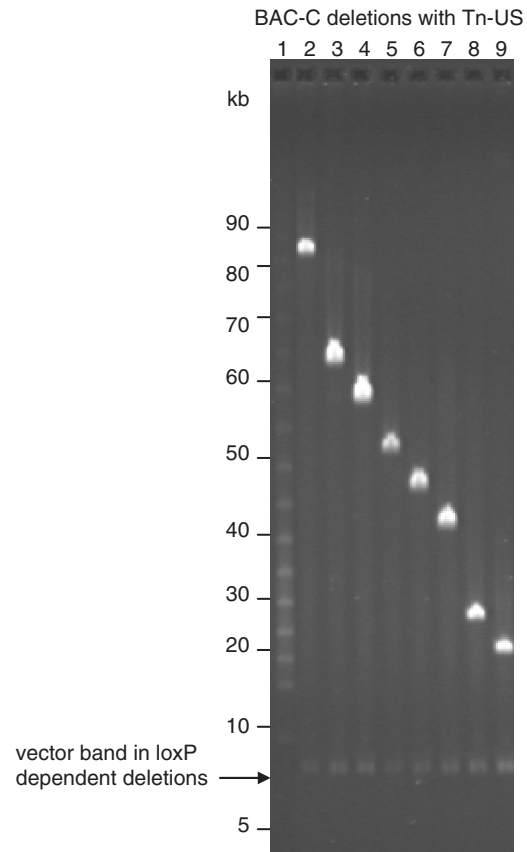

Figure 2. A FIGE analysis of DNA from BAC-C deletion clones generated with Tn-US. The DNA was digested with NotI enzyme before applying to gel. Lane 1 shows a 5-kb ladder.

was first identified: we generated a set of transposons that contained different lengths of this intron (0.8, 1.5 kb and the set of four constructs shown in Figure 4). These Tn-DNAs were injected into zebrafish embryos and GFP expression monitored. Thus, inclusion of fragments A or C (Figure 1B) did not express GFP, but fragment B resulted in notochord-specific GFP expression (Figure 4A and C).

All constructs contained the +0.147 to −0.75 kb (or +0.147 to −0.35 kb for Tn-3 and Tn-82) UE and varying lengths of intron 1 DNA flanking the GFP gene. A distinct pattern of GFP fluorescence in the notochord upon injecting linear transposon plasmid DNA into eggs was used as a preliminary assay for scoring the intron 1 deletion series (Figure 4A–D). It is known that the notochord is somewhat promiscuous to expressing a few other genes (43), but in our experience does not confer expression to any segment of DNA juxtaposed into the GFP cassette in the transposon plasmid. For example, segments A or C from intron 1 failed to induce expression of GFP (data not shown). The 1.2-kb piece within segment B (Tn-3 and Tn-82 in Figure 1C) was demonstrated to be both required and sufficient for the specific pattern of notochord expression (Figure 4). No other pattern of GFP expression was observed in any other tissue with the intron 1 deletion series of Tn-plasmids.
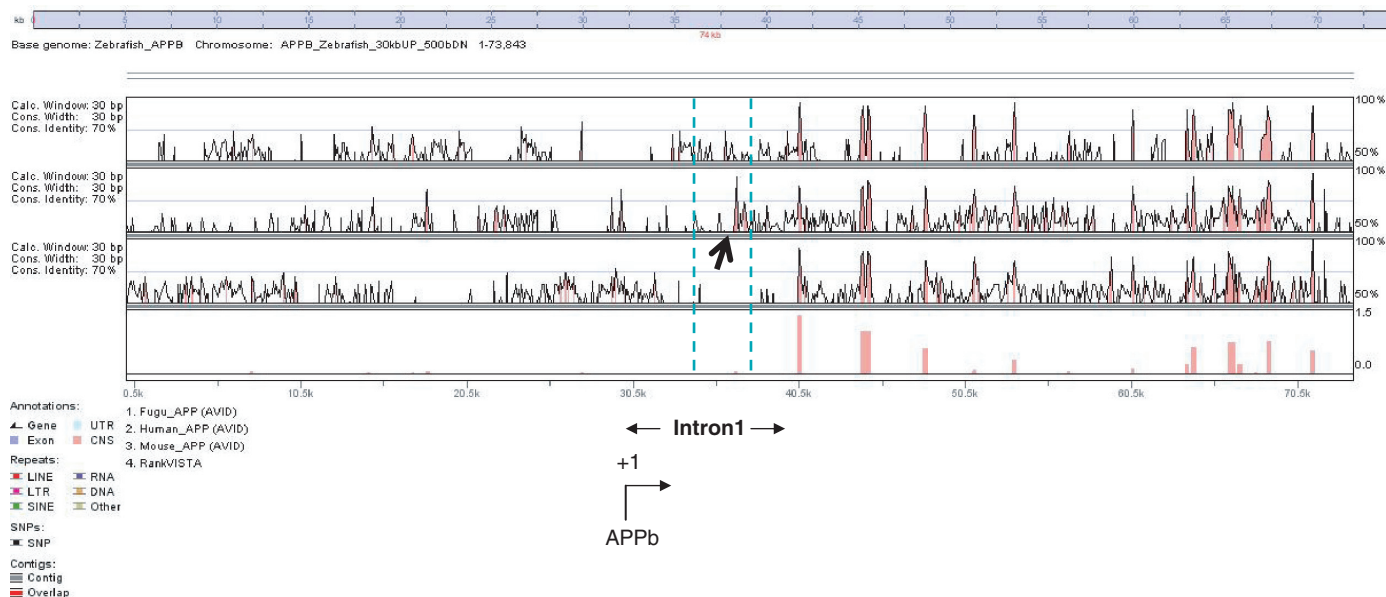
Figure 3. Comparative genome sequence analysis of the APPb gene region −30 kb upstream to +0.5 kb downstream with respect to the zebrafish APPb gene transcription start site. Each horizontal panel represents a pair-wise comparison of the vertebrate species with zebrafish. The first row of peaks show a comparison of Fugu with zebrafish, the second row human with zebrafish and the third row compares mouse with zebrafish. The location of the transcription start site is indicated as +1 and the intron 1 interval is marked. The region of interest within intron 1 is marked by the green vertical broken lines. Note the high degree of conservation between human and zebrafish in intron 1 marked by the thick black arrow within the green vertical lines. The orange vertical bar designating this is not prominent because this particular conservation is absent in mouse.

Semi-conserved element in APPb intron 1 behaves as a true enhancer

The role of orientation of intron 1 element (IE) in enhancing APPb gene expression was tested by inverting the 2.2-kb fragment B in the transposon. Fragment B activated expression in both orientations (indicated as Tn-13 and Tn-10) and produced identical patterns of GFP expression (see Figure 4A and C), indicating that the segment behaved as a true enhancer. Opposite orientations of this enhancer in traps, Tn-3 and Tn-82, also produced similar results, either in transposons (Figure 4B and D) or when integrated into BAC DNA (compare Δ78C and Δ75C, or Δ52C and Δ38C in Figures 6 and 7). No open reading frames could be identified in the 1.2-kb minimal enhancer. This intron 1 enhancer could not substitute for the +0.147 to −0.75 kb UE and was unable to express GFP by itself when placed downstream of the GFP reporter (data not shown).

The sequence of intron 1 enhancer DNA common to Tn-3 and Tn-82 (Figure 1C) maps to zebrafish chromosome 9 draft assembly ZFISH7:9:29144733-29145740, and is shown in Figure 5. A truncated version, with approximately 0.8 kb of this enhancer sequence (Tn-46, not shown), does not express GFP in the notochord as efficiently as the 1.2 kb intron 1 enhancer (data not shown).

Generating BAC deletions with enhancer trap transposons carrying ~1.2 kb intron 1 enhancer (IE) and 0.35 kb upstream sequence (UE) fused to GFP

Having rescued partial activity of the APPb promoter, we next used this minimal construct to identify additional sequences upstream of the gene that are essential for expression characteristic of endogenous APPb (22). Several deletion series were generated in BACs C and D with Tn-3 and Tn-82. A FIGE analysis of APPb BAC-C deletions generated with Tn-3 and Tn-82 is shown in Figure 6. Insert DNA in deletions range in size 30–100 kb. Three deletions made in BAC-D with Tn-82 were also used to complete the analysis. Locations of enhancer trap ends in the sequence contig surrounding APPb, mapped by BAC end sequencing (27), is shown to scale in Figure 7B.

Approximately 60 kb on either side of the APPb gene is devoid of other genes. There is an ASMT-like expressed sequence tag (EST) mapped ∼60 kb upstream of APPb. The next gene on the 5′ side is NCAM2, and is ∼160 kb away (Figure 7B).

Patterns of GFP expression from enhancer trap BACs carrying ~1 kb intron 1 (IE) and 0.35 kb upstream sequence (UE) fused to GFP

The enhancer trap retrofitted BAC DNAs were injected into zebrafish eggs without linearization at 0 h postfertilization. Embryos were scored between 48 and 72 h later using a fluorescence microscope. The pattern of expression of GFP fluorescence shown in Figure 7A and C, is representative of the clone and derived from at least three successful experiments. We consider an experiment successful when the survival of embryos is between 30% and 60% of those injected 48 h prior to scoring, and positive expression occurs in at least 20% of those survived. A typical injection used around 60–100 eggs. Although all of the
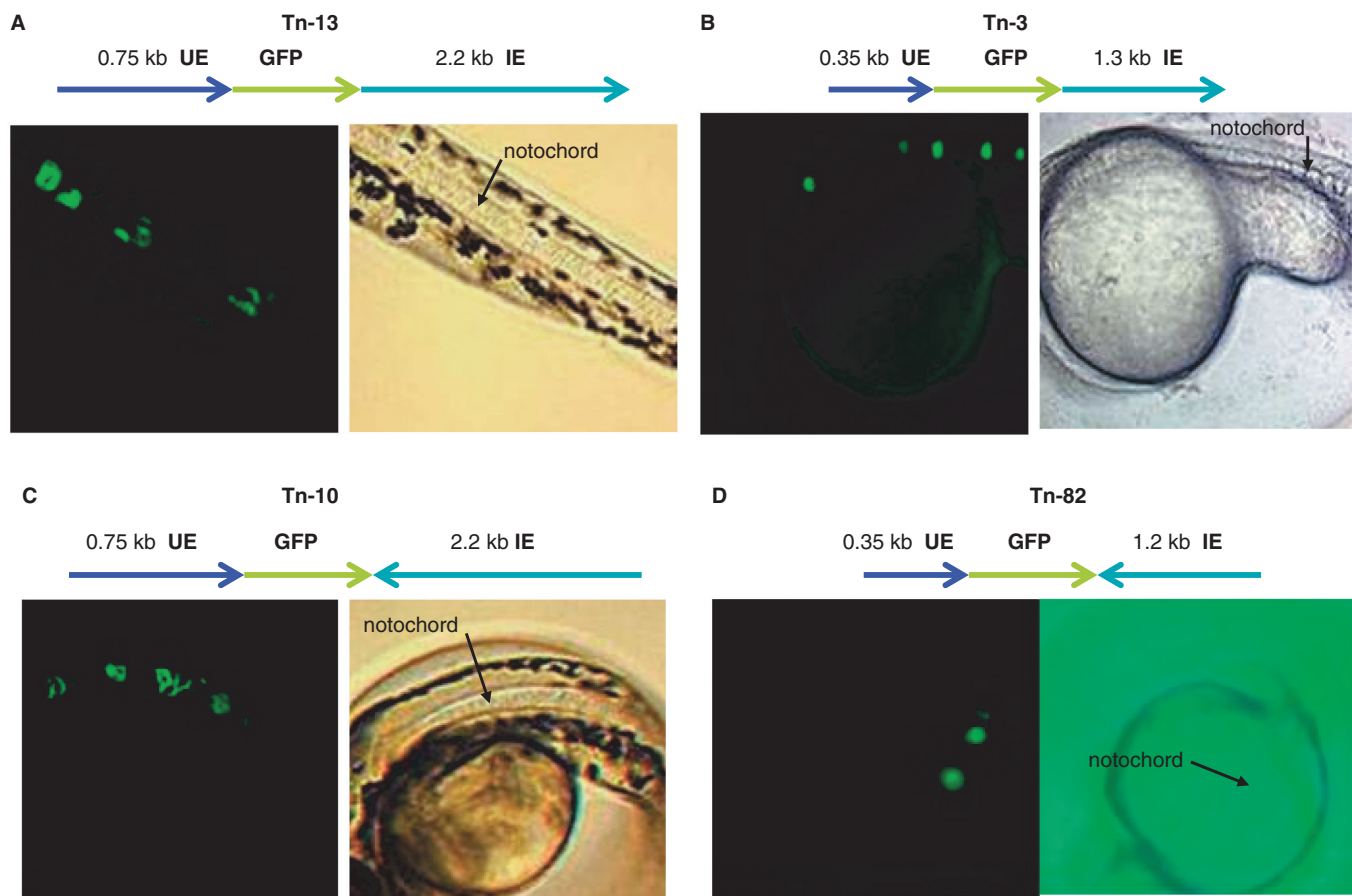
Figure 4. Patterns of GFP expression observed by injecting enhancer trap transposon plasmid. A schematic of the DNA construct injected is shown on top of each panel with the color coding and orientation of UE, GFP and IE identical to Figure 1C. Each pair comprises the fluorescence and phase contrast of the same animal. (**A–D**) display the result of injecting DNAs from Tn-13, Tn-3, Tn-10 and Tn-82, respectively. Location of notochord is marked by arrow. The embryos were between 24 and 60 h postfertilization when photographed. Note that expression is exclusive to the notochord.

deletion clones shown in Figure 7B, were analyzed for GFP expression, the data for only a subset of these clones are shown. Figure 7A shows the pattern of GFP expression for clones Δ74D, Δ70D and Δ94C. In sharp contrast to the fluorescence patterns generated with the enhancer trap transposon plasmids (Figure 4), distinct patterns of expression specific to neuronal cells was observed when injecting DNAs from this set of BACs. There is a unique network pattern of GFP fluorescence in the midbrain region, and this pattern persists till Δ84C, as seen in Figure 7A. Bright field images for some of these are shown in Supplementary Figure 1. Expression is also seen in neurons of the spinal chord extending throughout the length of the fish in this set (not shown). Fluorescence in neurons is indicated by the white arrowheads. The numbers in column of Figure 7D (left) represent injected fish with positive expression out of 100 surviving 72 hpf.

Expression specific to neurons was also observed with Δ80C through Δ75C (Figure 7A) but pattern was simpler, with fluorescence mostly along the spinal chord. BLAST of the end sequence of Δ75C puts its end to be about 28 kb upstream of the transcription start site of APPb.

DNA from BACs Δ72C, Δ52C, Δ38C and Δ29C produced no neuronal expression of GFP (Figure 7C). Instead, all of the fluorescence was restricted to just the notochord, and is reminiscent of the pattern observed earlier with injecting the enhancer trap transposon plasmids Tn-10, Tn-13, Tn-3 and Tn-82 (Figure 4). BLAST analysis places the ends of Δ52 and Δ38 to 46 kb and 60 kb, upstream of APPb.

Germline expression of enhancer trap BACs

Stable transgenic fish lines have been isolated from a few of the enhancer trap BACs shown in Figure 7B, and the results from these corroborate our findings from transient expression. Two independent lines of germline fish were isolated from each of the two BACs Δ94C and Δ84C. The F1s from different founders with the same BAC DNA vary somewhat in their expression patterns, although neural fluorescence is seen in all of them. Expression from two lines of Δ94C and Δ84C is displayed in Figures 8 and 9, respectively. The yolk sac has much higher levels of GFP fluorescence between the 48 and 72 hpf window of observation in all stable lines tested. Difficulty in imaging the neural fluorescence against this

Intron-1 Enhancer Sequence in CNE of APPb gene
Assembly: ZFISH7: chr 9: position 29144733-29145740

```
TAAAACTAGACCACAGCACAATCCCTTTCGGTCAGTATTGTTATTACATAAGACATC
AGCAATAAAGAAAAAGATTGACCCACTGTAATTGGTTTTCACATTAGTATTGATTCAC
GTAACACTACAGATGAATGGAATCAAGTTTGTGGGTCTTGGTACATGTGTGCTATAG
CTTTGCCATTGTTCCAAAACTGTATTTAGTTAATTTAATTCTTCAAAATTAATTAGTAA
GTCTTAAGCATGTCAAGTGCTTAAATCTGACCTCTTGTAAGTCATTATTTTCTTTCTG
ACTGGTTCAATAAAAAAGCCCCTTTAGGCTTTTTTCATTATAGGTGCCACTTGTAGG
TACTGTGCCAAAACAATCTTTGACAGTTTGCGCTGTCACCTCTTTACTGGTCACTAT
AACCACACACAAGTCATTTTGGCTGAAAGGTCCTTTGCTTTAGCTTGCTTGTAATTA
CTAAATCTGTTTTCAGTCAGCCTCCATCTGAAAAAGGCCTTTCTTTGTTTGACTGGG
AAATGTCACTGGAACATGTAAACACGGTCGGTTTTAAATTATGTTTTTGAGAGCTAC
AAAAGCTACTTCATTGATTTTGACAGCCTTCAAGATTAAGCCAGAGCTGAATGACTC
TTGAAAGCTGCACATTTCATTTTGATAGATGCATTAATGTGGATCCTGAGCTGCTAT
GACAACCTGACTGAGACACTATTTCGGGAGATTGATATTCGCAGATGATTAGCTTTC
ACCAATGGCTCAGTGGATGTCGGTGTTAAAGCATTTTGAGCACCATGGTAATTCTAT
CATAAATGTACATCTATAGCTTATGTGCAGCATACAGAATGATCTTACTACTGCAAT
GAGCTGTCGCACACATGGGTTTCCTAAAGAGCTAACTACTGTACTGCGGTCAGTGT
AATTACAGTCTTCTGTTTTTAAACTGCCATACAGCAGAGAGATGTTTTATTGCGCTTT
AACCACCAAGCTCTTTGATTAAAAGATAGATAGAAGGAAAAACAGATTGCCAGAGTC
TATC
```

Figure 5. Sequence of the intron 1 enhancer of zebrafish APPb gene. Only the sequence of 1008 bp common to the IEs of transposons Tn-3 and Tn-82 is shown here.
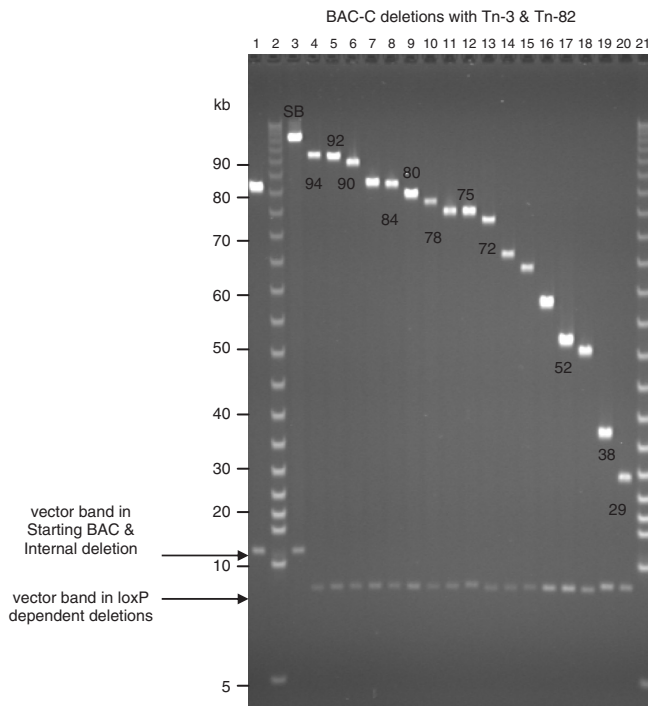
Figure 6. FIGE analysis of DNA isolated from enhancer trap-modified BAC deletions after Not I digestion. Deletions generated in BAC-C with Tn-3 (Δ75C, Δ38C and Δ29C) are shown in lanes 12, 19 and 20, respectively, and those with Tn-82 shown in lanes 4–11 and 13–18. All deletion clones Δ94C, through Δ29C are numbered according to the apparent sizes of their insert DNA on FIGE, and these are indicated directly above or below the insert DNA band. A loxP-independent internal deletion is shown in lane 1, and its vector band is identical to that in starting BAC-C shown in lane 3 as SB. Lanes 2 and 21 show a 5-kb ladder.

yolk sac background is probably exacerbated by the low copy numbers of BACs known to integrate into the germline compared to small plasmids. The problem is more severe in the head region apparently due to its proximity, and dissipates towards the tail. At still later times, pigmentation increases and interferes with imaging. The network

pattern of GFP fluorescence seen in Figure 9A–C for Δ84C, and Figure 8C for Δ94C, appears to originate from glial cells and oligodendrocytes thought to ensheathe axons (44). These, along with the fluorescence in neurons of the lateral line, are components of the peripheral nervous system (PNS). Fluorescence from neurons of the lateral line, is clearly visible below the stellate cells (marked by the successive pink arrows in Figures 8A–C and 9A–C), and appear similar to those observed in transient expression patterns of those BACs seen earlier in Figure 7A. Such expression patterns are not observed in embryos that do not carry the BAC DNA in their germline as seen in Figure 8D. This control embryo shown was obtained from parents that were positive for transient expression of GFP characteristic of the Δ94C expression pattern; but which failed to transmit the BAC DNA through the germline. Overall, the fraction of fish with transient expression that transmitted to the next generation was slightly below 1%. A large fraction of the transients died during the 2-week period following injection. No >7% of the transiently expressing embryos surviving till adulthood were found to transmit the DNA through the germline, i.e. their eggs were GFP positive. Additionally, over 90% of embryos from such germline founders were found to express GFP fluorescence. Thus, expression in neurons, from possibly both the central nervous system (CNS) and PNS, from both stable transgenic lines with BACs Δ94C and Δ84C substantiate some of our earlier findings with transient expression.

Context dependence of the intron 1 enhancer

Patterns of expression shown for Δ74D, Δ70D, Δ94C, Δ92C and Δ84C in Figure 7A, and expression in stable lines derived from Δ94C and Δ84C (shown in Figures 8 and 9), resemble the pattern of APPb expression reported earlier using *in situ* hybridization of APPb mRNA probes (22). One needs to take into account the differences in signal-to-noise resolution between a sandwich assay (22), and direct fluorescence analyzed here. Importantly, Δ74D through Δ75C did not express in the notochord (Figure 7C). Thus, including the genome context of APPb suppressed inappropriate expression in the notochord, and activated it specifically in cells where the endogenous APPb gene expresses. A different type of context dependence of GATA factor function has been reported recently, where different regulatory modules dictate its activity in hematopoietic versus endothelial cells (24). Interestingly, the context dependent APPb intron 1 enhancer sequence identified here also contains binding sites for members of the GATA factor family; with high scores for GATA-3 (see Supplementary Figure 2).

BioInformatic analysis of transcription factor-binding sites within intron 1 enhancer

BioInformatic analysis for transcription factor binding sites was conducted for the ~1 kb intron 1 enhancer DNA. The results are shown in Supplementary Figure 2. High scores for SOX 5 and GATA 3 factor binding sites were noted.
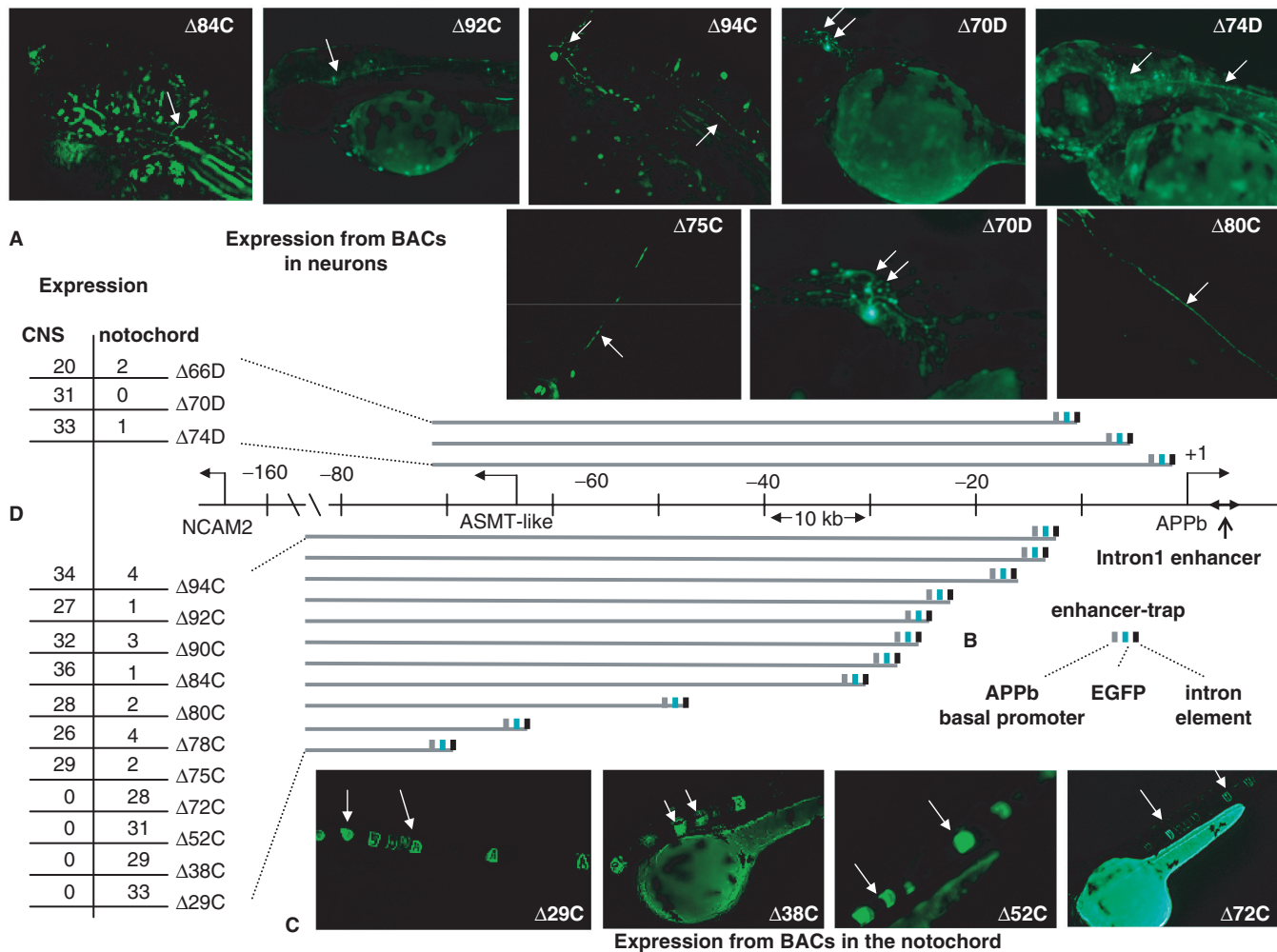
Figure 7. (**A**) Pattern of GFP fluorescence observed with injecting DNA from Enhancer trap retrofitted BACs: transient expression patterns produced by BACs Δ74D, Δ70D, Δ94C, Δ92C and Δ84C have an extensive network pattern of fluorescence in the mid- and hind-brain regions (indicated by white arrows) in addition to fluorescence in the spinal chord that runs throughout the length of the fish. An enlarged view of a section of the expression pattern with BAC Δ70D is shown immediately below the whole picture to emphasize the network pattern in the mid- and hind-brain regions. Only neural expression is observed with these BACs along with BACs Δ80C through Δ75C, although the pattern is much simpler in the latter set. Note the fluorescence in the middorsal ganglia extending all the way covering the entire length of the fish in each of these BACs. These are also indicated by the white arrows. There is no expression in the notochord in BACs Δ74D through Δ75C. (**B**) A schematic drawn to scale of BAC deletions Δ74D through Δ29C is shown against the sequence contig containing the APPb gene as assembled in version ZFISH7. The enhancer trap containing ends of deletions from BACs-C and -D are shown as colored triplets, and their locations were determined by sequencing with transposon end-based primers. The locations of APPb, ASMT-like EST and NCAM2 genes are indicated on the sequence contig. Also indicated by the double-edged arrow is the location of the enhancer in intron 1 of APPb. (**C**) Shows the strikingly different pattern of GFP expression observed with enhancer trap containing BACs Δ72C, Δ52C, Δ38C and Δ29C. Expression is exclusively in the notochord as shown by the white arrows. (**D**) A table of numbers showing the distribution of fish expressing either in neurons (CNS) or cells of the notochord for each enhancer trap BAC injected. These are out of 100 animals surviving till 72 hpf. The pictures here were taken between 60 and 72 hpf. Bright field images for several pictures shown here are included in Supplementary Figure 1.

DISCUSSION

Regulation of the APPb gene

Several earlier reports noted the absence of a consensus TATA box in the −30 region of the APP gene in higher vertebrates (15,18). Although the zebrafish APPb gene promoter remains uncharacterized, regulation of the gene from several other vertebrates had been studied extensively (15–21). As much as 8000 bp upstream of the transcription start site has been shown to contain regulatory activity in conventional CAT assays (18).

The 5′-untranslated region (UTR) has also been implicated to contain regulatory elements responsive to iron (45), interleukin-1 (46) and TGF-β (47). In order to preserve such regulation in germline transgenic fish derived from these enhancer trap-modified BACs, the UE in our enhancer trap transposons include the 147 bp of this UTR (Figure 1B).

Comparison of data in Figures 4 and 7–9, demonstrate that the intron 1 enhancer identified here directs specific expression of a GFP-reporter gene in two completely different tissues depending on whether APPb promoter
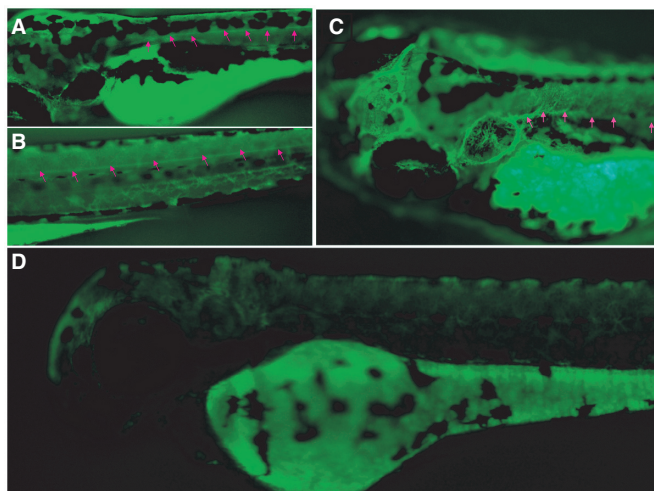
Figure 8. Expression in F1 animals derived from germline transgenic parents that were injected with DNA from BAC Δ94C are shown. Strong fluorescence from the yolk sac combined with low copy numbers of integrated BAC presents hurdles to data collection. (**A** and **B**) Show GFP fluorescence in the lateral line (indicated by the pink arrows) in the anterior and posterior portions of the same fish. (**C**) Shows the fluorescence in an embryo from a different germline parent. (**D**) Shows a control embryo that is GFP negative from a parent that was positive in transient expression of GFP. The overall fluorescence is low in the control compared to those of GFP positive embryos under identical conditions and, more importantly, never shows fluorescence in the lateral line. Pictures were recorded between 48 and 72 hpf.
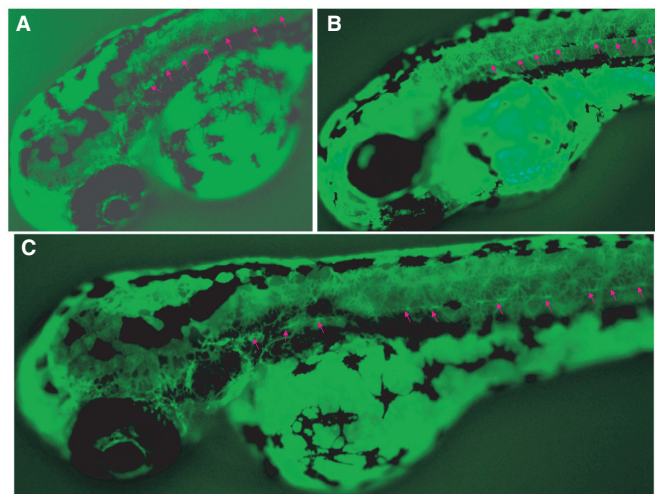
Figure 9. Expression in F1 animals derived from germline transgenic parents that were injected with DNA from BAC Δ84C is shown. (**A–C**) Shows GFP expression in the lateral line (indicated by the pink arrows), and possibly in the dorsal root ganglion, visible just beneath the extensive network pattern of fluorescence on the surface.

proximal or far upstream promoter elements of the gene are adjacent to it. We conclude that the enhancer can function specifically in nonneural tissue such as the notochord when used with promoter proximal elements within the $+0.147$ to -0.35 kb of sequence surrounding the transcription start site of APPb. However, this promiscuity disappears and its function becomes

exquisitely specific to enhancing expression in neurons when juxtaposed with additional promoter elements located farther upstream till about -28 kb of APPb. The data also suggest that some type of transcription repression activity resides around -28 kb that suppresses expression in the notochord, because simultaneous expression in the notochord and neural cells is not observed in the same fish using BAC deletions Δ74D through Δ75C (Figure 7A). Once this upstream DNA extending till -28 kb is deleted, as in BACs Δ72C, Δ52C, Δ38C and Δ29C, reappearance of the notochord expression pattern is observed (Figure 7C). Nevertheless, BACs Δ72C, Δ52C, Δ38C and Δ29C, serve as important controls to show that the exclusive neural pattern observed with BACs Δ74D through Δ75C is not merely a consequence of the enhancer trap being in the BAC environment as opposed to the small Tn-plasmid, but that far-upstream promoter elements located within the ~28 kb DNA upstream of the start site are required for neuron-specific expression. The strikingly different, yet specific, expression patterns elicited by this enhancer when working in or out of context of its own gene is unique to the best of our knowledge.

Tissue-specific expression of the APPb gene resembling its endogenous pattern (22) requires two separate, somewhat distant regulatory domains to cooperate in *cis* to confer tissue-specificity. The two domains of regulation are the ~1 kb of DNA within intron 1 (ZFISH7:9:29144733-29145740), and the region $+0.147$ to -28 kb upstream of the gene. As indicated above in this report, exclusion of the ~1 kb IE produced no expression of GFP: APPb BAC-C deleted from the wild-type loxP end of insert DNA with Tn-US (Figures 1B, C and 2) failed to express GFP in any tissue (data not shown). Absence of the $+0.147$ to -28 kb region on the other hand leads to expression in the inappropriate tissue, as seen with both the set of Tn-plasmids Tn-10, Tn-13, Tn-3 and Tn-82 (Figure 4), and BACs Δ72C, Δ52C, Δ38C and Δ29C (Figure 7C).

A search for transcription factor binding sites in the ~1 kb intron enhancer sequence using bioinformatic tools reveals a high likelihood for GATA 3 and SOX 5 recognition sites (Supplementary Figure 2). Although definitive identification of actual binding by these transcription factor complexes to this enhancer region awaits further studies, it is rather intriguing to note that GATA 3 and associated factors have been implicated in pathways guiding the aging process in *C. elegans* (48), and involved also in allergic inflammation (49). SOX 5 has been implicated, along with SOX 9, in the differentiation and establishment of several cell lineages including chondrocytes and glial cells of the nervous system in the spinal chord (50). Binding sites for SOX 9 and other SRY-related protein factors are also evident within the 1 kb intron enhancer sequence as indicated in Supplementary Figure 2.

Results from both transient expressions (Figure 7), and stable lines with BACs Δ84C and Δ94C (Figures 8 and 9), suggests that upstream regulation of the APPb gene is likely to extend much farther than the 8 kb limit identified for the primate APP earlier (18). Ends of BACs Δ94C

and Δ84C map to 13.6 and 24.1 kb upstream of the APPb transcription start site, respectively, and neural expression from these clones in stable lines therefore suggest that APPb-specific regulatory sequences exist at locations beyond −24 kb that, in conjunction with the 0.35 kb UE and 1.2 kb IE, can assemble productive transcription initiation complexes specifically in neurons to express GFP. The earlier determination that 8 kb is the limit of regulation by upstream sequences in primate APP (18) appears unlikely; as we note that the zebrafish genome is compressed ∼4-fold with respect to that of rhesus monkey.

The simpler pattern of expression with BAC deletions Δ80C through Δ75C (Figure 7A) might indicate that residual promoter elements required for neural cell-specific expression are available till about 28 kb upstream, and these might prove sufficient to produce partial expression patterns by recruitment of factors through protein–protein interactions similar to those described earlier (51,52). Importantly, there is no expression in the notochord from these BACs.

The upstream regulatory regions identified here probably pertain to APPb only, and are unadulterated by those from adjoining genes: the 200 kb upstream of APPb has only two other genes, one ASMT-like EST annotated some 60 kb away and another, NCAM2, located ∼160 kb from APPb (Figure 7B). Sequences downstream of APPb intron 1 are deleted in all BACs during retrofitting with enhancer traps.

Highlights of the methodology

The methodology described here uses BACs, and therefore has all of the advantages associated with their use compared to small plasmid constructs (53–60). Unlike the other BAC recombineering approaches; however, our strategy is 'nontargeted', and produces a large set of DNA constructs rapidly to test for function of potential regulatory elements without having to select sequences for modifications. It is also not affected by repetition of sequences elsewhere in the chromosome. Subtractive analysis from one or both ends (25), to decipher the role of individual regulatory elements in a cluster, is also likely to be easier using the enhancer trap BAC approach described here.

In situations where function is conserved without sequence similarity (8), choosing the correct sequence to test might be a hit-or-miss phenomenon with a targeted approach. Our methodology should prove most helpful here since it acts more like a 'mine-sweeper' moving along the BAC DNA; capable of identifying all regulatory sequences upstream of a gene in an unbiased manner. Functional comparisons between individual enhancer trap BACs are thus more meaningful as the modifications in each remain constant. While the Tol2 system can score functional CNEs using zebrafish rapidly (10); subtractive analysis for exploring mechanisms of combinatorial interplay of multiple elements regulating a gene, similar to that described here, would be tedious. The BAC enhancer trap technology has three additional features that might also prove beneficial: (i) allows sampling much larger DNA, and consequently multiple discontinuous regulatory domains simultaneously, (ii) the context of regulatory DNA with respect to the gene and chromosome is preserved and (iii) can be used with BACs in established libraries from a wide variety of organisms, and tested in several species. Although the methodology does not allow generation of internal deletions, truncations from the end opposite to the enhancer trap can be made with a lox511 transposon to explore functions of candidate regulatory regions in a limited way. Sequences bending DNA (61–63), or phasing nucleosomes and other transcription factors (64–66) are left unaltered using BACs compared to characteristics of the gene region found endogenously. Bringing exogenous pieces of DNA together to create artificial joints in small plasmids to *trans*-activate reporter genes do not adequately address the endogenous role of the regulatory sequence, and this is avoided using BACs. DNA structure surrounding regulatory factor binding sites have evolved over long periods, and these are also left unaltered here. We note there are 52 sites with six or more A-residues, known to cause bends in unpackaged DNA (62), in the 28-kb upstream regulatory sequence identified here in APPb.

Trapping gene regulatory elements by random insertion of a transposon containing reporter gene with minimal promoter into the genomes of Drosophila (67,68), and other vertebrates (11,69–73) has been very successful. These approaches insert the enhancer trap, i.e. p-element, retroviral vector or transposable element, directly into the genome of the organism, with mapping of the sites of integration done subsequently. BACs in contrast have already been mapped on the chromosome.

The BAC enhancer trap methodology is amenable to manipulation along several lines: GFP reporter in the transposon can be replaced by clinically relevant mutant cDNAs of the gene, such as the Swedish mutant form of human APP (30), to create animal models for screening small molecules to generate drug leads. Instead, toxigenes replacing the GFP reporter could produce tools for cell killing in specific tissue in transgenic animals.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Ahituv,N., Prabhakar,S., Poulin,F., Rubin,E.M. and Couronne,O. (2005) Mapping cis-regulatory domains in the human genome using multi-species conservation of synteny. *Hum. Mol. Genet.*, **14**, 3057–3063.
2. Ahituv,N., Rubin,E.M. and Nobrega,M.A. (2004) Exploiting human–fish genome comparisons for deciphering gene regulation. *Hum. Mol. Genet.*, **13**, R261–R266.
3. Dermitzakis,E.T., Reymond,A. and Antonarakis,S.E. (2005) Conserved non-genic sequences–an unexpected feature of mammalian genomes. *Nat. Rev. Genet.*, **6**, 151–157.
4. Shin,J.T., Priest,J.R., Ovcharenko,I., Ronco,A., Moore,R.K., Burns,C.G. and MacRae,C.A. (2005) Human-zebrafish non-coding conserved elements act *in vivo* to regulate transcription. *Nucleic Acids Res.*, **33**, 5437–5445.
5. Woolfe,A., Goodson,M., Goode,D.K., Snell,P., McEwen,G.K., Vavouri,T., Smith,S.F., North,P., Callaway,H., Kelly,K. *et al.* (2005) Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol.*, **3**, e7.
6. McEwen,G.K., Woolfe,A., Goode,D., Vavouri,T., Callaway,H. and Elgar,G. (2006) Ancient duplicated conserved non-coding elements in vertebrates: a genomic and functional analysis. *Genome Res.*, **16**, 451–465.
7. Xie,X., Kamal,M. and Lander,E.S. (2006) A family of conserved noncoding elements derived from an ancient transposable element. *Proc. Natl Acad. Sci. USA*, **103**, 11659–11664.
8. Fisher,S., Grice,E.A., Vinton,R.M., Bessling,S.L. and McCallion,A.S. (2006) Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science*, **14**, 276–279.
9. Pennacchio,L.A., Ahituv,N., Moses,A.M., Prabhakar,S., Nobrega,M.A., Shoukry,M., Minovitsky,S., Dubchak,I., Holt,A., Lewis,K.D. *et al.* (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, **444**, 499–502.
10. Fisher,S., Grice,E.A., Vinton,R.M., Bessling,S.L., Urasaki,A., Kawakami,K. and McCallion,A.S. (2006) Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat. Protocols*, **1**, 1297–1305.
11. Ellingsen,S., Laplante,M.A., Konig,M., Kikuta,H., Furmanek,T., Hoivik,E.A. and Becker,T.S. (2005) Large-scale enhancer detection in the zebrafish genome. *Development*, **132**, 3799–3811.
12. Shizuya,H., Birren,B., Kim,U.J., Mancino,V., Slepak,T., Tachiiri,Y. and Simon,M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl Acad. Sci. USA*, **89**, 8794–8797.
13. Ioannou,P.A., Amemiya,C.T., Garnes,J., Kroisel,P.M., Shizuya,H., Chen,C., Batzer,M.A. and de Jong,P.J. (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.*, **6**, 84–89.
14. Osoegawa,K., Mammoser,A.G., Wu,C., Frengen,E., Zeng,C., Catanese,J.J. and de Jong,P.J. (2001) A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res.*, **11**, 483–496.
15. Salbaum,J.M., Weidemann,A., Lemaire,H.G., Masters,C.L. and Beyreuther,K. (1988) The promoter of Alzheimer's disease amyloid A4 precursor gene. *EMBO J.*, **7**, 2807–2813.
16. Yoshikai,S.I., Sasaki,H., Dohura,K., Fururya,H. and Sakaki,Y. (1990) Genomic organization of the human amyloid beta-protein precursor gene. *Gene*, **87**, 257–263.
17. Lahiri,D.K. and Robakis,N.K. (1991) The promoter activity of the gene encoding Alzheimer beta-amyloid precursor protein (APP) is regulated by two blocks of upstream sequences. *Brain Res. Mol. Brain Res.*, **9**, 253–257.
18. Song,W. and Lahiri,D.K. (1998) Functional identification of the promoter of the gene encoding the Rhesus monkey beta-amyloid precursor protein. *Gene*, **217**, 165–176.
19. Lahiri,D.K., Song,W. and Ge,Y.W. (2000) Analysis of the 5′-flanking region of the beta-amyloid precursor protein gene that contributes to increased promoter activity in differentiated neuronal cells. *Brain Res. Mol. Brain Res.*, **77**, 185–198.
20. Richardson,J.C., Kendal,C.E., Anderson,R., Priest,F., Gower,E., Soden,P., Gray,R., Topps,S., Howlett,D.R., Lavender,D. *et al.* (2003) Ultrastructural and behavioural changes precede amyloid deposition in a transgenic model of Alzheimer's disease. *Neuroscience*, **122**, 213–228.
21. Lahiri,D.K., Ge,Y.W. and Maloney,B. (2005) Characterization of the APP proximal promoter and 5′-untranslated regions: identification of cell-type specific domains and implications in APP gene expression and Alzheimer's disease. *FASEB J.*, **19**, 653–665.
22. Musa,A., Lehrach,H. and Russo,V.A. (2001) Distinct expression patterns of two zebrafish homologues of the human APP gene during embryonic development. *Dev. Genes Evol.*, **211**, 563–567.
23. Erman,B. and Sen,R. (1996) Context dependent transactivation domains activate the immunoglobulin mu heavy chain gene enhancer. *EMBO J.*, **15**, 4665–4675.
24. Wozniak,R.J., Boyer,M.E., Grass,J.A., Lee,Y. and Bresnick,E.H. (2007) Context dependent GATA factor function: combinatorial requirements for transcriptional control in hematopoietic and endothelial cells. *J. Biol Chem.*, **282**, 14665–14674.
25. Shakes,L.A., Garland,D.M., Srivastava,D.K., Harewood,K.R. and Chatterjee,P.K. (2005) Minimal cross-recombination between wild type and loxP511 sites *in vivo* facilitates truncating both ends of large DNA inserts in pBACe3.6 and related vectors. *Nucleic Acids Res.*, **33**, e118.
26. Chatterjee,P.K. (2004) Retrofitting BACs and PACs with LoxP transposons to generate nested deletions. In Zhao,S. and Stodolsky,M. (eds), *Bacterial Artificial Chromosomes*, Vol. 1. The Humana Press Inc., Totowa, NJ, USA, pp. 231–241.
27. Chatterjee,P.K., Yarnall,D.P., Haneline,S.A., Godlevski,M.M., Thornber,S.J., Robinson,P.S., Davies,H.E., White,N.J., Riley,J.H. and Shepherd,N.S. (1999) Direct sequencing of bacterial and P1 artificial chromosome nested-deletions for identifying position-specific single nucleotide polymorphisms. *Proc. Natl Acad. Sci. USA*, **96**, 13276–13281.
28. Chatterjee,P.K. and Baker,J.C. Jr. (2004) Preparing nested deletions template DNA for field inversion gel electrophoresis analyses and position-specific end sequencing with transposon primers. In Zhao,S. and Stodolsky,M. (eds), *Bacterial Artificial Chromosomes*, Vol. 1. The Humana Press Inc., Totowa, NJ, USA, pp. 243–254.
29. Dillen,K. and Annaert,W. (2006) A two decade contribution of molecular cell biology to the centennial of Alzheimer's disease: are we progressing toward therapy? *Int. Rev. Cytol.*, **254**, 215–300.
30. Reaume,A.G., Howland,D.S., Trusko,S.P., Savage,M.J., Lang,D.M., Greenberg,B.D., Siman,R. and Scott,R.W. (1996) Enhanced amyloidogenic processing of the beta-amyloid precursor protein in gene-targeted mice bearing the Swedish familial Alzheimer's disease mutations and a 'humanized' Abeta sequence. *J. Biol. Chem.*, **271**, 23380–23388.
31. Gilmore,R.C., Baker,J. Jr., Dempsey,S., Marchan,R., Corprew,R.N.L. Jr., Maeda,N., Smithies,O., Byrd,G., Bukoski,R.D., Harewood,K.R. *et al.* (2001) Using PAC nested-deletions to order contigs and microsatellite markers at the high repetitive sequence containing Npr3 gene locus. *Gene*, **275**, 65–72.
32. Brake,R.L., Chatterjee,P.K., Kees,U.R. and Watt,P.M. (2004) The functional mapping of long-range transcription control elements of the *HOX11* proto-oncogene. *Biochem. Biophys. Res. Com.*, **313**, 327–335.
33. Chi,X., Chatterjee,P.K., Wilson,W. III, Zhang,S.-X., DeMayo,F. and Schwartz,R.J. (2005) Complex cardiac Nkx2-5 gene expression activated by noggin sensitive enhancers followed by chamber specific modules. *Proc. Natl Acad. Sci. USA*, **102**, 13490–13495.

34. Banerji,J., Rusconi,S. and Schaffner,W. (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, **27**, 299–308.

35. Humphries,R.K., Ley,T., Turner,P., Moulton,A.D. and Nienhuis,A.W. (1982) Differences in human alpha-, beta- and delta-globin gene expression in monkey kidney cells. *Cell*, **30**, 173–183.

36. Berg,P.E., Yu,J.K., Popovic,Z., Schumperli,D., Johansen,H., Rosenberg,M. and Anderson,W.F. (1983) Differential activation of the mouse beta-globin promoter by enhancers. *Mol. Cell. Biol.*, **3**, 1246–1254.

37. Ameres,S.L., Drueppel,L., Pfleiderer,K., Schmidt,A., Hillen,W. and Berens,C. (2005) Inducible DNA-loop formation blocks transcriptional activation by an SV40 enhancer. *EMBO J.*, **24**, 358–367.

38. Zheng,R., Shen,R., Goodman,O.B. Jr. and Nanus,D.M. (2006) Multiple androgen response elements cooperate in androgen regulated activity of the type 1 neutral endopeptidase promoter. *Mol. Cell. Endocrinol.*, **259**, 10–21.

39. Guo,X., Evans,T.R., Somanath,S., Armesilla,A.L., Darling,J.L., Schatzlein,A., Cassid,J. and Wang,W. (2007) In vitro evaluation of cancer-specific NF-kappaB-CEA enhancer-promoter system for 5-fluorouracil prodrug gene therapy in colon cancer cell lines. *Br. J. Cancer*, **97**, 745–754.

40. Pheasant,M. and Mattick,J.S. (2007) Raising the estimate of functional human sequences. *Genome Res.*, **17**, 1245–1253.

41. Cooper,G.M. and Brown,C.D. (2008) Qualifying the relationship between sequence conservation and molecular function. *Genome Res.*, **18**, 201–205.

42. Way,J.C. and Kleckner,N. (1985) Transposition of plasmid-borne Tn*10* elements does not exhibit simple length-dependence. *Genetics*, **111**, 705–713.

43. Müller,F., Chang,B., Albert,S., Fischer,N., Tora,L. and Strähle,U. (1999) Intronic enhancers control expression of zebrafish sonic hedgehog in floor plate and notochord. *Development*, **126**, 2103–2116.

44. Pogoda,H.M., Sternheim,N., Lyons,D.A., Diamond,B., Hawkins,T.A., Woods,I.G., Bhatt,D.H., Franzini-Armstrong,C., Dominguez,C., Arana,N. *et al.* (2006) A genetic screen identifies genes essential for development of myelinated axons in zebrafish. *Dev. Biol.*, **298**, 118–131.

45. Rogers,J.T., Randall,J.D., Cahill,C.M., Eder,P.S., Huang,X., Gunshin,H., Leiter,L., McPhee,J., Sarang,S.S., Utsuki,T. *et al.* (2002) An iron-responsive element type II in the 5′-untranslated region of the Alzheimer's amyloid precursor protein transcript. *J. Biol. Chem.*, **277**, 518–528.

46. Shaw,K.T., Utsuki,T., Rogers,J., Yu,Q.S., Sambamurti,K., Brossi,A., Ge,Y.W., Lahiri,D.K. and Greig,N.H. (2001) Phenserine regulates translation of beta-amyloid precursor protein mRNA by a putative interleukin-1 responsive element, a target for drug development. *Proc. Natl Acad. Sci. USA*, **98**, 7605–7610.

47. Maloney,B., Ge,Y.-W., Greig,N. and Lahiri,D.K. (2004) Presence of a 'CAGA box' in the APP gene unique to amyloid plaque-forming species and absent in all APP-1/2 genes: implications in Alzheimer's disease. *FASEB J.*, **18**, 1288–1290.

48. Budovskaya,Y.V., Wu,K., Southworth,L.K., Jiang,M., Tedesco,P., Johnson,T.E. and Kim,S.K. (2008) An elt-3/elt-5/elt-6 GATA transcription circuit guides aging in *C. elegans*. *Cell*, **134**, 291–303.

49. Barnes,P.J. (2008) Role of GATA-3 in Allergic Diseases. *Curr. Mol. Med.*, **8**, 330–334.

50. Hattori,T., Coustry,F., Stephens,S., Eberspaecher,H., Takigawa,M., Yasuda,H. and de Crombrugghe,B. (2008) Transcriptional regulation of chondrogenesis by coactivator Tip60 via chromatin association with Sox9 and Sox5. *Nucleic Acids Res.*, **36**, 3011–3024.

51. Kaufmann,J. and Smale,S.T. (1994) Direct recognition of initiator elements by a component of the transcription factor IID complex. *Genes Dev.*, **8**, 821–829.

52. Nikolajczyk,B.S., Dang,W. and Sen,R. (1999) Mechanisms of mu enhancer regulation in B lymphocytes. *Cold Spring Harb. Symp. Quant. Biol.*, **64**, 99–107.

53. Giraldo,P. and Montoliu,L. (2001) Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Res.*, **2**, 83–103.

54. Yang,X.W., Model,P. and Heintz,N. (1997) Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat. Biotechnol.*, **9**, 859–865.

55. Zhang,Y., Buchholz,F., Muyrers,J.P. and Stewart,A.F. (1998) A new logic for DNA engineering using recombination in *Escherichia coli*. *Nat. Genet.*, **20**, 123–128.

56. Jessen,J.R., Meng,A., McFarlane,R.J., Paw,B.H., Zon,L.I., Smith,G.R. and Lin,S. (1998) Modification of bacterial artificial chromosomes through chi-stimulated homologous recombination and its application in zebrafish transgenesis. *Proc. Natl Acad. Sci. USA*, **95**, 5121–5126.

57. Muyrers,J.P., Zhang,Y., Testa,G. and Stewart,A.F. (1999) Rapid modification of bacterial artificial chromosomes by ET recombination. *Nucleic Acids Res.*, **27**, 1555–1557.

58. Gong,S., Yang,X.W., Li,C. and Heintz,N. (2002) Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6Kgamma origin of replication. *Genome Res.*, **12**, 1992–1998.

59. Warming,S., Costantino,N., Court,D.L., Jenkins,N.A. and Copeland,N.G. (2005) Simple and highly efficient BAC recombineering using galK selection. *Nucleic Acids Res.*, **33**, e36.

60. Yang,Z., Jiang,H., Chachainasakul,T., Gong,S., Yang,X.W., Heintz,N. and Lin,S. (2006) Modified bacterial artificial chromosomes for zebrafish transgenesis. *Methods*, **39**, 183–188.

61. Fried,M.G. and Crothers,D.M. (1983) CAP and RNA polymerase interactions with the lac promoter: binding stoichiometry and long range effects. *Nucleic Acids Res.*, **11**, 141–158.

62. Ulanovsky,L., Bodner,M., Trifonov,E.N. and Choder,M. (1986) Curved DNA: design, synthesis, and circularization. *Proc. Natl Acad. Sci. USA*, **83**, 862–866.

63. Gartenberg,M.R., Ampe,C., Steitz,T.A. and Crothers,D.M. (1990) Molecular characterization of the GCN4-DNA complex. *Proc. Natl Acad. Sci. USA*, **87**, 6034–6038.

64. Hebbar,P.B. and Archer,T.K. (2007) Chromatin-dependent cooperativity between site-specific transcription factors *in vivo*. *J. Biol. Chem.*, **282**, 8284–8291.

65. Ganapathi,M., Singh,G.P., Sandhu,K.S., Brahmachari,S.K. and Brahmachari,V. (2007) A whole genome analysis of 5′ regulatory regions of human genes for putative cis-acting modulators of nucleosome positioning. *Gene*, **391**, 242–251.

66. Hardy,S. and Shenk,T. (1989) E2F from adenovirus-infected cells binds cooperatively to DNA containing two properly oriented and spaced recognition sites. *Mol. Cell Biol.*, **9**, 4495–4506.

67. O'Kane,C.J. and Gehring,W.J. (1987) Detection in situ of genomic regulatory elements in Drosophila. *Proc. Natl Acad. Sci. USA*, **84**, 9123–9127.

68. Wilson,C., Pearson,R.K., Bellen,H.J., O'Kane,C.J., Grossniklaus,U. and Gehring,W.J. (1989) P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. *Genes Dev.*, **3**, 1301–1313.

69. Balciunas,D., Davidson,A.E., Sivasubbu,S., Hermanson,S.B., Welle,Z. and Ekker,S.C. (2004) Enhancer trapping in zebrafish using the sleeping beauty transposon. *BMC Genom.*, **5**, 62.

70. Korn,R., Schoor,M., Neuhaus,H., Henseling,U., Soininen,R., Zachgo,J. and Gossler,A. (1992) Enhancer trap integrations in mouse embryonic stem cells give rise to staining patterns in chimaeric embryos with a high frequency and detect endogenous genes. *Mech. Dev.*, **39**, 95–109.

71. Burns,J.C., Friedmann,T., Driever,W., Burrascano,M. and Yee,J.K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc. Natl Acad. Sci. USA*, **90**, 8033–8037.

72. Lin,S., Gaiano,N., Culp,P., Burns,J.C., Friedmann,T., Yee,J.K. and Hopkins,N. (1994) Integration and germ-line transmission of a pseudotyped retroviral vector in zebrafish. *Science*, **265**, 666–669.

73. Grabher,C., Henrich,T., Sasado,T., Arenz,A., Wittbrodt,J. and Furutani-Seiki,M. (2003) Transposon-mediated enhancer trapping in medaka. *Gene*, **322**, 57–66.

74. Chatterjee,P.K., Mukherjee,S., Shakes,L.A., Wilson,W. III, Harewood,K.R. and Byrd,G. (2004) Selecting transpositions of a markerless transposon using phage P1 headful packaging: new transposons for functionally mapping long range regulatory sequences in BACs. *Anal. Biochem.*, **335**, 305–315.