Volume 10 Number 9 1982 Nucleic Acids Research

Isolation and characterisation of the *Xenopus laevis* albumin genes: loss of 74K albumin gene sequences by library amplification

Felicity E.B.May*, Rudolf Weber+ and Bruce R.Westley*

Abteilung für Zell- und Entwicklungsbiologie, Baltzerstrasse 4, CH-3012 Bern, Switzerland

Received 25 March 1982; Accepted 15 April 1982

ABSTRACT

The blood of the frog X.laevis contains 2 albumins of 68,000 and 74,000 daltons which are encoded in the liver by two related mRNAs. When an amplified X.laevis DNA library was screened with cloned albumin cDNA only 68,000 dalton albumin gene sequences were isolated. Hybridisation of the albumin cDNA to Southern-blots of Eco R1 digested X.laevis DNA showed that the sequences present in the recombinants did not account for all the fragments which hybridised on the Southern-blots. This indicated that 74K albumin gene sequences exist but that they are not present in the amplified DNA library. A X.laevis genomic library was therefore constructed and screened for albumin genes without amplification. Both 68K and 74K albumin gene sequences were isolated. Recombinants containing 74K albumin gene sequences grew extremely poorly and this probably explains why the 74K albumin sequences were not isolated from the amplified library. Characterisation of the cloned DNA indicates that there is one sequence coding for the 68K albumin but two different sequences coding for the 74K albumin.

INTRODUCTION

Closely related pairs of genes code for the vitellogenins (1), adult and larval α and β globins (2) and ribosomal proteins (3) in the frog X.laevis. This recent data is consistent with an earlier suggestion (4) that the X.laevis genome is tetraploid as a result of a genome duplication 30 million years ago.

The blood of X.laevis contains a major 74,000 dalton and a minor 68,000 dalton albumin. We have recently shown (5) that the 2 albumins are encoded by two mRNAs which show about the same degree of mismatch as the pairs of vitellogenin and globin mRNAs suggesting that the albumin gene was also duplicated as part of the genome duplication. Since the duplication event, however, the albumin genes have diverged so that they now code for 2 proteins of very different molecular weights which are expressed at different levels.

To determine the features of the albumin genes which are responsible for these differences we attempted to isolate the 2 genes from a X.laevis DNA library. Although 68K albumin gene sequences could be isolated from an amplified X.laevis DNA library, 74K albumin gene sequences could only be isolated

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

0305-1048/82/1009-2791$2.00/0

from an unamplified library. Analysis of the albumin sequences obtained from the unamplified library suggests that there is one 68K but two distinct 74K albumin gene sequences in X.laevis.

METHODS

Preparation of Charon 4A Arms

The left and right arms of Charon 4A DNA were prepared after digestion with Eco RI essentially as described by Maniatis (6).

Preparation of X.laevis DNA Fragments

DNA was prepared from the liver of a single mature X.laevis female. Aliquots (50 µg/ml) were digested with a wide range of Eco RI concentrations (0.1-4 U per µg of DNA) in 100 mM Tris (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 50 µg/ml BSA at 37°C for 1h. The extent of digestion was monitored by electrophoresis on a 0.5% agarose gel; the aliquots were combined, brought to 0.1 M NaCl and extracted with phenol. The DNA was precipitated with ethanol and redissolved at 500 µg/ml in 10 mM Tris (pH 7.5), 0.1 mM EDTA. EDTA was added (to 10 mM) and the DNA incubated at 65°C for 20 min before sedimentation through 10-35% linear sucrose gradients as described (6). 500 µl fractions were collected, the size of the DNA measured as above and those fractions which contained DNA between 12-23 kb were dialysed against 10 mM Tris (pH 8.0), 5 mM NaCl, pooled and precipitated with ethanol.

Ligation of DNA and In Vitro Packaging of the Recombinant DNA into Phage Particles

1.25 µg of X.laevis DNA was ligated with 3.75 µg of purified Charon 4A arms by incubation with 0.8 units of T_4-DNA ligase in 20 mM Tris (pH 7.5), 10 mM $MgCl_2$, 0.6 mM ATP, 50 µg/ml BSA at 8°C for 8h.

In vitro packaging extracts were prepared from bacterial strains BHB 2688 and BHB 2690 as recommended by B. Hohn (personal communication). The cloning efficiency was 1.6×10^6 pfu per µg of DNA which allowed 1 haploid X.laevis genome equivalent of X.laevis DNA to be cloned from less than 0.2 µg of size selected X.laevis DNA. This efficiency is equivalent to 5.3×10^5 pfu of Charon 4A DNA. 1×10^4 pfu were obtained per µg of purified Charon 4A arms alone (background 2%).

Screening

Phages were screened for albumin gene sequences essentially as described by Wahli and Dawid (7). Nick-translated isolated fragments of cloned albumin cDNA (5) were used as the hybridisation probes. The fragments used are shown in Fig. 1 and were Hpa II fragments of pcXa 1 and pcXa 11 (containing the

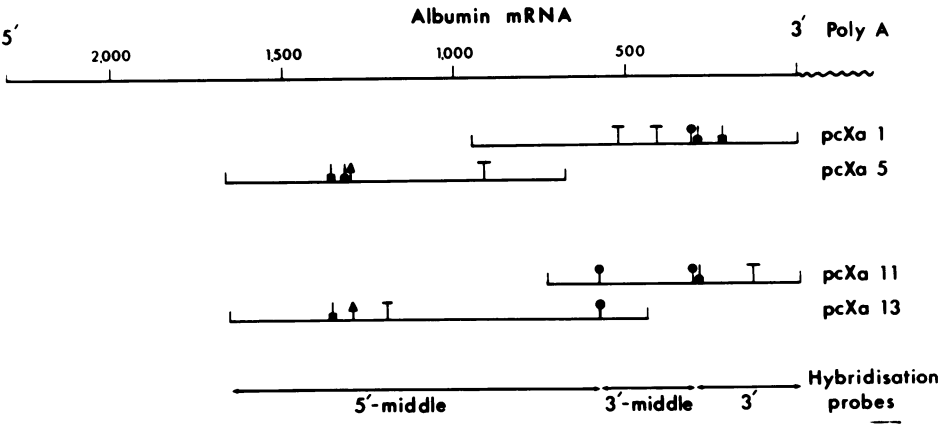


Figure 1. Restriction maps of cloned X.laevis albumin cDNAs used to screen the DNA libraries and to map the genomic clones.
The library was screened with nick-translated isolated fragments of cloned cDNA complementary to 1,700 nucleotides of the 74K (pcXa 1, pcXa 5) and 68K (pcXa 11, pcXa 13) albumin mRNAs. Albumin mRNA is drawn schematically above the clones. The genomic clones were mapped by hybridisation to a nick-translated 3' (Pst 1 fragment of pcXa 11), 3'-middle (Pst 1 fragment of pcXa 11), 5'-middle (Pst 1 fragment of pcXa 13) or kinase-labelled poly (A)+ RNA from male X.laevis liver. ▲, Hae III; ▼, Pst 1; ▲, Eco R1; T Alu 1; ↑, Hind III. Artificial Pst 1 sites are at the ends of the cloned cDNA.

entire cDNA insert plus a small amount of flanking vector DNA), the Pst 1 fragment of pcXa 5 containing the entire cDNA insert and the long Pst 1 fragment of pcXa 13. Positive plaque regions were purified through several rounds and DNA was prepared from albumin recombinants essentially as described (7). For recombinants which grew poorly, fresh high titre lysates were prepared to inoculate the large cultures and the ratio of recombinants to bacteria in the inoculum was increased 10-20 fold. Lysis took more than 24h.

Washing and Melting Curves

A single plaque was picked from a pure plate of each recombinant and the phage suspended in 100 µl of TMG (50 mM Tris pH 7.5, 10 mM $MgSO_4$, 0.01% gelatin). After dilution, about 300 pfu were plated out onto 8 cm plates and two nitrocellulose filter replicas were made of each plate. The first set of filters were hybridised to a mixture of pcXa 1 and pcXa 5 and the second set to pcXa 11 and pcXa 13, at 0.5×10^6 cpm/filter in 1 ml of hybridisation solution at 37°C for 14h. The filters were washed at room temperature; twice for 1h in 50% formamide, 2x SSC, 0.1% SDS and then twice for 30 min in 2x SSC, 0.1% SDS, dried and exposed to preflashed X-ray film. For the washing curves, 12 discs (6 mm diameter) were stamped out from positive regions of each fil-

ter, marked and the Cherenkov radiation was counted in 4x SSC; triplicate filters were then washed for five times 1h in 50% formamide, 0.1% SDS plus: 2x SSC; 0.5x SSC; 0.1x SSC or 0.05x SSC at 37°C. The discs were recounted to determine the percentage of radioactivity eluted. For the melting curves, individual discs were incubated for 10 min at increasing temperatures in 1 ml of 6 mM Tris (pH 8.0), 30 mM NaCl, 1 mM EDTA (temperature from 40°C to 85°C in 5° increments). The amount of hybridised radioactivity melted at each temperature was determined and expressed as a percentage of the total.

^{32}P-Labelled Probes

Cloned cDNA was labelled by nick-translation to specific activities of about 10^8 cpm/µg (8). Cytoplasmic RNA was isolated from the livers of male X.laevis and fractionated on poly(U)-Sepharose as described previously (9) and then end-labelled using γ-^{32}P-ATP and polynucleotide kinase.

Rapid Phage DNA Isolation

Phage lysates (10 ml) were prepared from randomly selected phage plaques and the phages pelleted by centrifugation. DNA was prepared from the phages as described previously (10) except that after the ethanol precipitation, the DNA was extracted twice with phenol, once with ether, reprecipitated and then treated with RNAse (1 µg/ml) in 100 µl of 10 mM Tris (pH 7.5), 1 mM EDTA. 5 µl was then digested for 14h with 10 U of Eco R1 as described above.

Restriction Analysis

DNA was digested as recommended by the enzyme supplier and reactions were stopped by addition of 1/4 volume 10% ficoll, 50 mM EDTA, bromophenol blue and 5x buffer (1x buffer: 36 mM Tris, 30 mM NaH_2PO_4, 0.5 mM EDTA). After electrophoresis, DNA was stained with ethidium bromide, photographed and then transferred to nitrocellulose (11). Prehybridisation, hybridisation and washing of filters was essentially as described (12). Cloned cDNA fragments were isolated and nick-translated as described (5).

RESULTS

Isolation of Genomic Albumin Sequences from an Amplified X.laevis DNA Library

An amplified X.laevis DNA library (7) was screened for albumin sequences using cDNA fragments from 4 clones which contain sequences complementary to the most 3' 1,700 nucleotides of the closely related 74K and 68K albumin mRNAs (see Methods and Fig. 1). 1.5×10^6 recombinant phages were screened (equivalent to 7 haploid X.laevis genomes) and 12 recombinants containing albumin sequences were isolated. The sequences contained within these recombinants were characterised by restriction mapping and by determining the Tm

of hybrids formed with 74K and 68K albumin cDNA. This analysis showed that the 12 recombinants contained overlapping DNA fragments covering a total of 24 kb of the 68K albumin gene and flanking DNA (from approximately the 3rd Eco R1 site to 3 kb 3' of the 11th Eco R1 site; Fig. 7). No recombinants containing the 74K albumin gene were isolated from the amplified library.

Analysis of Albumin Sequences Present in Uncloned X.laevis DNA

To verify that a 74K albumin gene is present in X.laevis DNA, 74K albumin cDNA was hybridised to Southern-blots of Eco R1 digested X.laevis DNA which had been prepared from the liver of a single animal. Under the nonstringent hybridisation conditions used in this experiment the 74K albumin cDNA hybridises equally well to 68K and 74K albumin sequences. Fig. 2, track A shows the hybridisation pattern obtained with the cDNA of pcXa 5. Five detectable fragments of 6.0, 5.4, 5.1, 3.6 and 2.7 kb hybridised whereas this cDNA fragment hybridised to only 3 of these fragments (5.1, 3.6 and 2.7 kb) in the cloned 68K albumin gene.

Track B shows the hybridisation pattern obtained with the large Pst 1 fragment of pcXa 1. Three fragments of 5.4, 5.2 and 2.7 kb hybridised whereas this cDNA hybridised to only one of these fragments (2.7 kb) in the cloned 68K albumin gene. These experiments showed clearly that a minimum of 3 genomic Eco R1 fragments which contain albumin sequences (5.2, 5.4 and 6.0 kb) had not been isolated from the amplified X.laevis DNA library. It was concluded that these 3 fragments probably contain 74K albumin gene sequences, especially as they were detected in the DNA of a single animal whereas the amplified library had been constructed from the DNA of several individuals.

Construction and Screening of an Unamplified X.laevis DNA Library

A second DNA library was then constructed from the same DNA as was used for the experiments shown in Fig. 2. This library was screened for albumin sequences without amplification to increase the probability that the 74K albumin gene would be isolated.

To ensure that the length of the cloned DNA fragments varied over the expected range and contained different numbers of Eco R1 fragments, the DNA from 10 independent phages was prepared as described in the Methods, digested with Eco R1 and then electrophoresed through a 1% agarose gel (Fig. 3). The X.laevis DNA inserted into these 10 phages was between 10.03 and 19.4 kb long (average: 14.3) and was contained in 1-4 Eco R1 fragments (average: 2.5). This indicates that the library should contain any sequence except those contained in Eco R1 fragments longer than about 22 kb and should therefore contain the 74K albumin fragments detected in the Southern hybridisation exper-

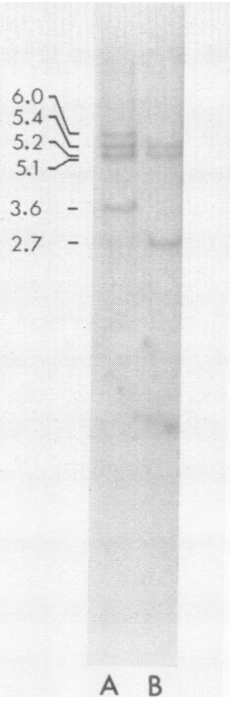


Figure 2. Hybridisation of albumin cDNA fragments to Eco R1 digested X.laevis liver DNA.

5 µg of the same DNA as was later used to prepare the X.laevis DNA library was digested with Eco R1, electrophresed on a 1% agarose gel, transferred to nitrocellulose and hybridised with the entire cDNA insert of pcXa 5 (isolated after Pst 1 digestion), (A) or the long Pst 1 fragment of pcXa 1, (B) which had been nick-translated to a specific activity of around 10^8 cpm/µg.

These two probes also hybridise to a fragment of 0.55 kb in cloned 68K albumin DNA but hybridisation to this fragment is too weak to be detected. The 5.1 kb fragment is also marked 5.2 kb because subsequent experiments (see text relating to Figures 6 and 7) showed this band to be the result of hybridisation to a 5.1 kb 68K albumin gene fragment and a 5.2 kb 74K gene fragment. Both these fragments hybridised in track A whereas only the 5.2 kb gene fragment hybridised in track B.

iments. 1.5×10^6 recombinant phages (about 7 X.laevis genomes) were screened for albumin sequences using the same fragments of X.laevis albumin cDNA as had been used to screen the amplified library.

In this first screening, 13 recombinant phages containing albumin sequences were identified. Three different types of hybridisation signal were obtained. 10 recombinants gave strong hybridisation signals (intense black spot on autoradiogram of about 6 mm diammeter, Fig. 4A), 2 gave strong hybridisa-

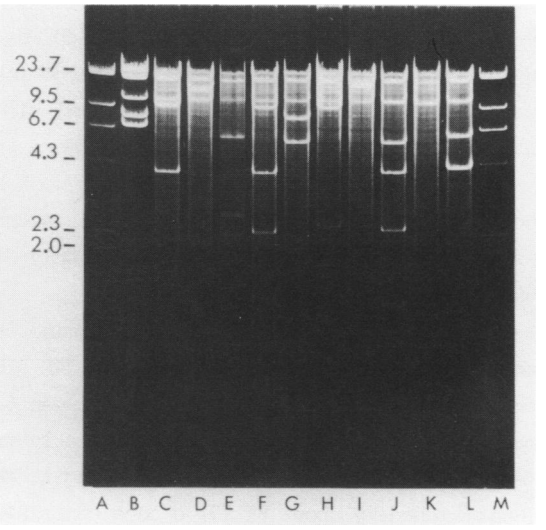


Figure 3. Agarose gel of Eco Rl digested DNA from 10 randomly selected recombinant phages.

Tracks C-L, Eco Rl digested DNA prepared from a 10 ml culture of 1 recombinant phage; tracks A and M, Hind 111 digested λ DNA; track B, Eco Rl digested Charon 4A DNA. The sizes of the marker fragments in tracks A and M are shown on the left in kb.

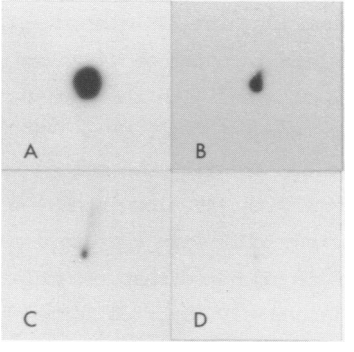


Figure 4. Autoradiographic signals obtained from 4 albumin recombinants during screening.

A, λX68a 206; B, λX74a 111; C, λX68a 209 and D, λX74a 102. Nitrocellulose filter replicas of plates containing these recombinants were treated identically, hybridised under the same conditions and exposed to X-ray film for the same length of time as described in the Methods.

tion signals but the spot on the autoradiogram was only 1.3 mm diammeter (Fig. 4B) and one gave a weak hybridisation signal which was also very small (Fig. 4C). During purification of the 13 phages the recombinants which had given the strongest hybridisation signals grew as normal sized plaques. The recombinant which had given the hybridisation signal shown in Fig. 4C continued to grow as small but visible plaques and to give a weak hybridisation signal whereas the two recombinants which had given the signals shown in Fig. 4B grew as plaques which were sometimes invisible and only detectable by their hybridisation signal.

We then used a rapid and convenient method to determine whether any of the 13 purified recombinants contained sequences from the 74K albumin gene. Washing and melting curves were made of hybrids formed between 68K or 74K albumin cDNA and recombinant phage DNA which was absorbed directly from the plaque onto nitrocellulose filters. As the cDNAs corresponding to the 68K and 74K albumin mRNAs are mismatched by about 8% (5) clones containing 74K albumin gene sequences could be identified by the Tm and salt sensitivity of the hybrids.

Typical washing and melting curves obtained with 2 recombinants (λX74a 111 and λX68a 208) are shown in Fig. 5. Fig. 5A shows that elution of 68K albumin cDNA from a filter containing λX68a 208 required a lower salt concentration than was required to elute 74K albumin cDNA. This indicates that the hybrids between λX68a 208 DNA and 68K albumin cDNA are more stable than hybrids formed with the 74K albumin cDNA and indicates that this phage contains 68K albumin gene sequences. When this experiment was repeated using recombinant λX74a 111 (Fig. 5B) the hybrids with 74K albumin cDNA were more stable which indicated that this recombinant contained 74K albumin gene sequences.

Typical melting curves obtained with the same two recombinants (washing curves; Fig. 5A and B) are shown in Fig. 5C and D. The hybrids formed between recombinant λX68a 208 and 68K albumin cDNA melted with a Tm of 66°C (o-o, Fig. 5C) while recombinants with 74K albumin cDNA melted at 62°C (●-●, Fig. 5C). Hybrids with 68K albumin cDNA were therefore more stable than hybrids With 74K albumin cDNA which suggested that recombinant λX68a 208 contains 68K albumin gene sequences. Conversely, hybrids formed between recombinant λX74a 111 and 68K albumin cDNA melted at 62°C (o-o, Fig. 5D) while hybrids with 74K albumin cDNA melted at 66.5°C (●-●, Fig. 5D) indicating that this clone contains 74K albumin gene sequences. In all cases, the assignment of the recombinant phage was the same from the washing and melting curves.

Only two of the 13 recombinants contained 74K albumin gene sequences. Sig-

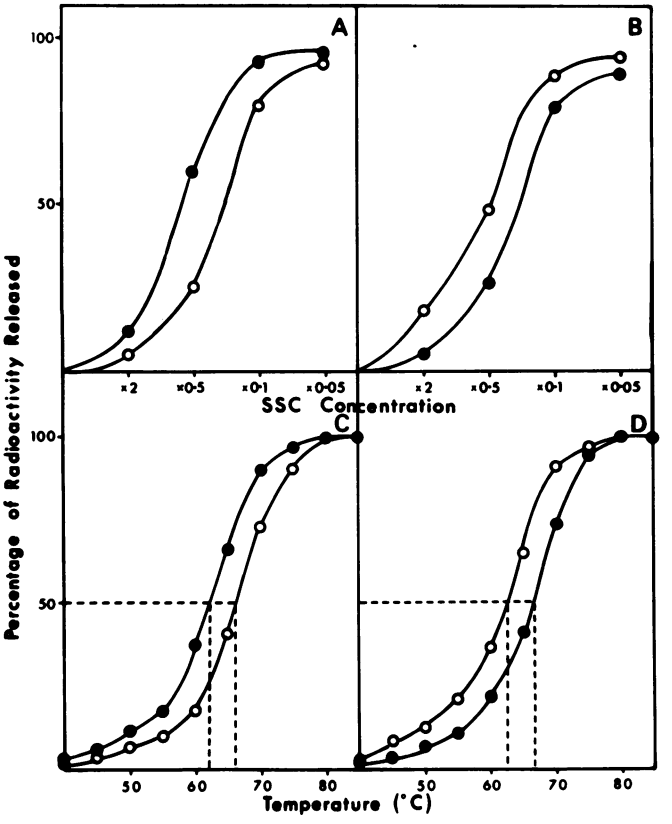


Figure 5. Washing and melting curves of hybrids formed between 68K or 74K albumin cDNA and DNA from phage plaques immobilised on nitrocellulose.

Nick-translated 74K (●–●) or 68K (o–o) albumin cDNA was hybridised to nitrocellulose filters containing λX74a 111 (A and C) or λX68a 206 (B and D) DNA which had been lifted directly from phage plaques onto nitrocellulose. Small circles cut from the filters were counted and then either washed at different SSC concentrations (A and B) or incubated at increasing temperatures (C and D) as described in the Methods. The percentage of the total radioactivity removed from the filter at each salt concentration or temperature is shown.

nificantly, these two recombinants had always grown as small plaques. These experiments, however, established that it was possible to clone 74K albumin gene sequences. Probably, we had not been able to isolate recombinant phages containing these sequences from an amplified library because they grow extremely poorly and would be strongly selected against during amplification.

To be more certain of isolating the entire 74K albumin gene, we screened a

further $7x10^5$ phages. We also attempted to isolate recombinants from areas of the original plates which had given weak hybridisation signals (previously scored as background) as it seemed likely that recombinants containing 74K albumin gene sequences would produce weak hybridisation signals due to the small size of the plaques. Six more albumin recombinants were isolated from the $7x10^5$ phages. Of these, 3 which grew as large plaques and one as small plaques all gave strong hybridisation signals and contained 68K albumin gene sequences. Two gave very weak hybridisation signals (Fig. 4D) which were derived from small plaques and contained 74K albumin gene sequences. Two further albumin recombinants which had given "background" hybridisation signals during screening of the first $1.5x10^6$ recombinants were isolated. One grew as very small plaques and contained 74K albumin gene sequences. The other which contained 68K albumin gene sequences gave slightly larger plaques and was later shown to contain only a short stretch of coding sequence which accounted for the weak hybridisation signal.

Characterisation of Albumin Recombinants

The 16 recombinants containing 68K albumin gene sequences and the 5 containing 74K albumin gene sequences were then mapped using single and double digests with 8 restriction enzymes. The restriction fragments were analysed on 0.8 and 1.5% agarose gels and hybridised sequentially with probes corresponding to various regions (see Fig. 1) of the albumin mRNAs after transfer to nitrocellulose filters. This analysis showed that 4 of the 5 74K albumin recombinants and 9 of the 16 68K recombinants contain different sequences (five were duplicated and one triplicated). The ethidium bromide stained Eco Rl digested DNA of the 13 different recombinants and the pattern of hybridisation with three different probes is shown in Fig. 6 and the restriction maps of all these clones is shown in Fig. 7.

The organisation of the various Eco Rl fragments could almost be deduced from the series of 9 overlapping clones. The clones contained a total of 35 kb of X.laevis DNA of which 13.3 kb, contained within 5 Eco Rl fragments (5.1, 3.6, 2.7, 2.2 and 0.55 kb, Fig. 6B), hybridised to kinase labelled RNA. 5 kb of DNA in the 5' flanking region and 17.5 kb in the 3' flanking region are contained within the clones and the gene is almost six times the length of the mRNA. The two largest Pst 1 fragments of pcXa 11 (3' and 3'-middle cDNA probes) hybridised to two Eco Rl fragments (2.7 and 2.2 kb, Fig. 6D). Individually, they each hybridised to one fragment (see Fig. 7) showing that the Eco Rl site in the cDNA clones which is very close to the Pst 1 site in pcXa 11 may be aligned with the Eco Rl site between the 2.2 and 2.7 kb gen-

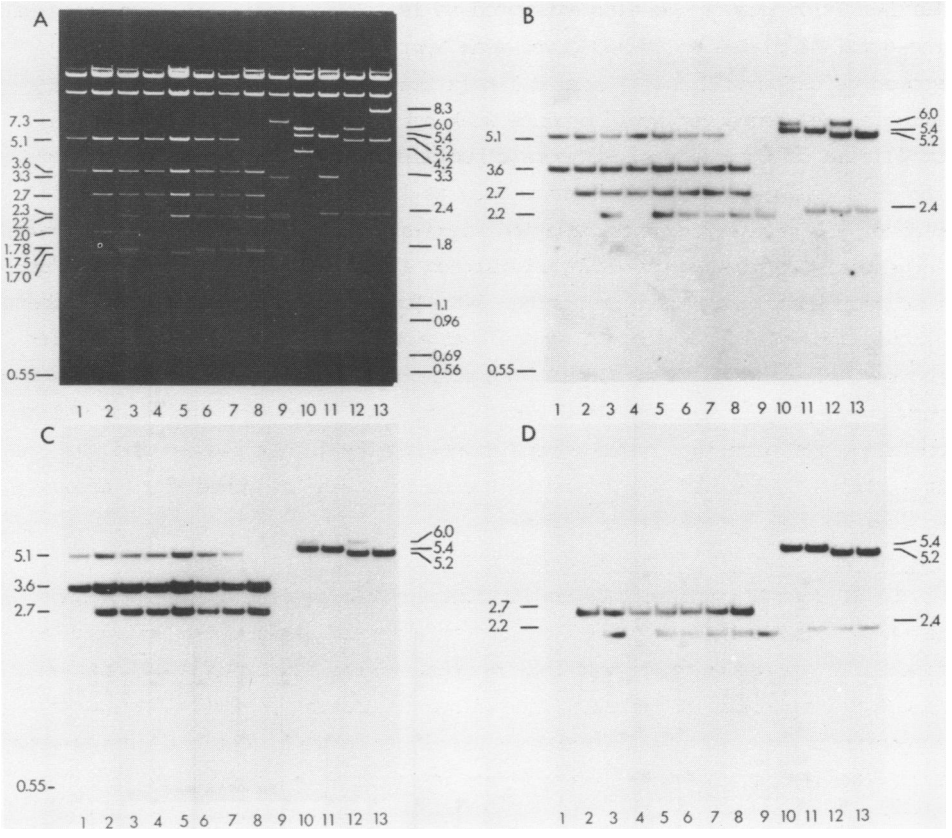
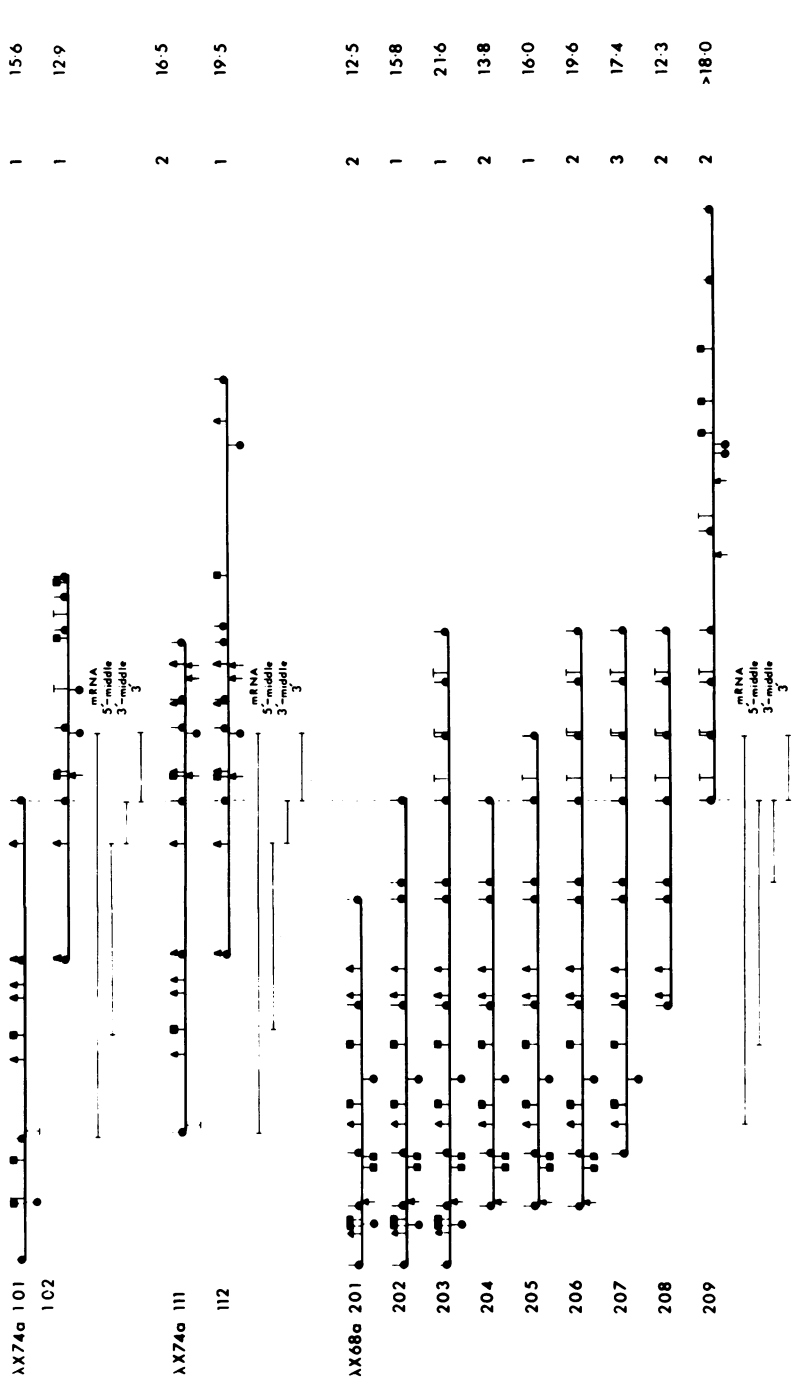


Figure 6. Hybridisation of albumin cDNA to Eco R1 digested cloned Genomic albumin DNA.

A, ethidium bromide stained 1% agarose gel of Eco R1 digested DNA prepared from the albumin genomic clones. Tracks 1-9, λX68a 201-209; tracks 10 and 11, λX74a 101 and 102 and tracks 12 and 13, λX74a 111 and 112. The size of the Eco R1 fragments within the 68K albumin clones are shown on the left and those of the 74K albumin clones on the right. The two bands at the top of the gel are the left and right arms of Charon 4A. The DNA was transferred to nitrocellulose and the autoradiographs are shown (B-D) after hybridisation with kinase-labelled poly (A)+ RNA from male X.laevis liver (B), nick-translated 5'-middle probe (C) or a mixture of nick-translated 3'-middle and 3' probes (D). These are described in Fig. 1. The sizes of the fragments hybridising are shown at the sides of the figure in kb.

mic fragments. No polymorphism was detected in any restriction site in the 9 68K albumin clones.

Restriction fragments of the 74K albumin recombinants were mapped using the 68K albumin cDNA fragments shown in Fig. 1 as the closely related 68K and 74K


Number isolated
Insert length (kb)
0 1 5 kb
λX74α 101
15·6
1
102
12·9
1
mRNA
5'-middle
3'-middle
3'
λX74α 111
16·5
2
112
19·5
1
λX68α 201
12·5
2
202
15·8
1
203
21·6
1
204
13·8
2
205
16·0
1
206
19·6
2
207
17·4
3
208
12·3
2
209
>18·0
2

albumin sequences cross-hybridise (Ref. 5 and Fig. 5). This showed that two distinct sequences of 23 kb and 25 kb of X.laevis DNA were present in the 4 different 74K albumin clones (Fig. 7). The length of DNA hybridising to kinase labelled RNA was about 13.5 kb for both sequences. The two 74K sequences had very different sites in the 3' flanking DNA. All restriction sites within the 74K gene were common to both sequences (Fig. 7) but there was a difference of 200 bp in their lengths. This difference is shown in Fig. 6 where the respective Eco R1 fragments in the middle of the 74K albumin genes (compare clone λX74a 111, track 12, and λX74a 101, track 10) are 5.2 and 5.4 kb long.

Fig. 6 shows that the mixed 3' and 3'-middle cDNA probe hybridised to 2 Eco R1 fragments in each 74K albumin sequence (2.4 and 5.2 kb, 2.4 and 5.4 kb). Individually these probes each hybridised to one fragment (Fig. 7). Thus the two 74K and one 68K genomic albumin sequences may be aligned on this common Eco R1 site (marked by a vertical dashed line in Fig. 7). There is also a Hind 111 site very close to an Eco R1 site in both 74K albumin gene sequences which probably corresponds to the Hind 111 site adjacent to the Eco R1 site in the 74K albumin cDNA about 1000 nucleotides from the 5' end of the 74K albumin mRNA sequence. This Hind 111 site presumably alignes with one of the two Hind 111 sites which are about 5 kb from the 5' end of the 68K albumin gene as the 68K albumin cDNA contains a Hind 111 site in the same position as the 74K albumin cDNA. In general however, although the lengths of the 74K and 68K albumin genes are extremely similar, their restriction maps are very different and there are very few common restriction sites.

Figure 7. Restriction maps of genomic albumin clones.
The 74K (λX74a 101-λX74a 112) and 68K (λX68a 201-λX68a 209) albumin clones have been aligned on the Eco R1 site (vertical dashed line) which is present in 74K and 68K albumin cDNA clones and lies at the border of the 3' and 3'-middle hybridisation probes (see Fig. 1). The number of independent clones isolated which contained the same sequence of X.laevis DNA and the total length of X.laevis DNA contained within the clones is shown on the right. The thin lines under the clones show the DNA which hybridises to the 3', 3'-middle, 5'-middle and RNA probes described in Fig. 1. ▲, Eco R1; ●, Bam H1; ▼, Bgl 11; ■, Xho 1; ▼, Kpn 1; ↑, Hind 111; I, Sac 1 and I, Sal 1.
We are not sure whether there are 1 or 2 Bgl 11 sites in the 2.0 kb Eco R1 fragment of λX68a 201-203, 1 or 2 Xho 1 sites in the 1.7 kb Eco R1 fragment of λX68a 201-206 or 1 or 2 Bam H1 sites in the >7 kb fragment of λX68a 209. In each case two sites have been drawn on the map. The relative orientation of the two smallest Hind 111 fragments at the 3' end of the 6.0 kb fragment of λX74a 101 and 111 is unknown. The >7 kb Eco R1 fragment of λX68a 209 was often heterogeneous in length, this heterogeneity was localised within the 1.8 kb Bgl 11 fragment and is thought to result from deletions during phage growth.

DISCUSSION

In this study we describe the isolation and characterisation of 3 distinct X.laevis genomic albumin sequences from an unamplified DNA library constructed from the liver DNA of a single animal. Comparison of the cloned and uncloned Eco R1 fragments which hybridise to cloned albumin cDNA suggests that we have isolated all genomic albumin sequences.

The identity of the genomic albumin sequences was established from the stability of hybrids formed with the previously characterised (5) 74K and 68K albumin cDNAs. This method, which is rapid and does not require the preparation of phage DNA, should be useful for the preliminary characterisation of recombinants containing sequences from other families of related genes.

One group of recombinants formed more stable hybrids with 68K than 74K albumin cDNA and is therefore presumed to contain 68K albumin gene sequences. The other two groups of recombinants formed more stable hybrids with 74K albumin cDNA and are therefore presumed to contain 74K albumin gene sequences.

The 68K and both 74K albumin gene sequences are all about 13.5 kb long which is almost six times longer than the albumin mRNA and is similar to the length of the rat and mouse albumin genes (16 and 17). The restriction maps of the 74K and 68K albumin gene sequences were very different probably because there are few common sites even within the 74K and 68K coding sequences (5) and because more than 80% of the genes are composed of intervening sequences which are thought to diverge faster than coding sequences (12).

No differences in any restriction sites analysed were detected in the 34 kb of DNA containing the 68K albumin gene. In addition, the restriction map of the 22 kb of DNA covering this gene which was isolated from the amplified library constructed from the DNA of several individuals (7) was the same. Therefore, no restriction site polymorphisms have been detected in or around the 68K albumin gene.

The discovery of 2 different 74K albumin gene sequences was completely unexpected and explained the complex patterns of hybridisation obtained when even small albumin cDNA fragments are hybridised to Southern-blots of uncloned DNA digested with various restriction enzymes. Both the 74K genomic albumin sequences formed equally well matched hybrids with 74K albumin cDNA and the restriction sites within the parts of the sequences which hybridised to albumin cDNA were very similar. There is, however, a difference in length of 200 bp located within a 5.4/5.2 kb Eco R1 restriction fragment and in several Hpa II and Hha I restriction sites (unpublished data). In contrast to the high degree of similarity within the sequences hybridising to albumin cDNA, there

was no discernible homology 3' of these sequences.

It is presently unclear whether the 2 74K genomic albumin sequences represent 2 polymorphic forms of the 74K albumin gene or 2 very similar 74K albumin genes. If the two 74K albumin sequences represent 2 distinct genes then their close similarity suggests that they are the result of an extremely recent duplication or an older duplication since which the similarity of the sequences has been maintained. The presence of 2 expressed 74K but one 68K albumin gene could account for the 2-fold difference in 74K and 68K albumin mRNA levels in X.laevis liver.

The albumin recombinants isolated from the unamplified library showed systematic differences in strength of hybridisation signals and plaque size. The strength of the hybridisation signal was related to the amount of coding sequence contained in the recombinant. Thus, the plaques of all but one recombinant containing 68K albumin sequences gave strong hybridisation signals because they contained most of the albumin coding sequence. The 68K albumin recombinant which gave a very weak signal only contained the extreme 3' end of the gene (λX68a 209)

Our results indicate that there are two different reasons why phages containing X.laevis DNA grow poorly. The plaques of all but two of the 68K albumin recombinants were of normal size. Of these two, λX68a 203 probably grows poorly because the DNA insert is extremely long (21.6 kb which is close to the maximum for Charon 4A) as it is all contained in other recombinants which grow normally. In contrast, the poor growth of λX68a 209 is probably due to a sequence within the X.laevis DNA. In DNA prepared from both isolates of this recombinant the largest Eco R1 fragment (>7 kb) was always heterogeneous in length, indicating that a sequence within this fragment which may have an inhibitory effect on phage growth was frequently deleted. The same is probably true for the 5 74K albumin recombinants because they all grow poorly although they do not contain especially long inserts of X.laevis DNA (12.9-19.5 kb).

Gene sequences are often isolated in unexpected frequencies from DNA libraries (eg: 16 and 17) and it has proved impossible to clone certain gene sequences using this approach (18). The observations that recombinants which contain 74K albumin gene sequences grow badly and are not present in an amplified library suggest that the above difficulties may result from library amplification rather than an inability of the DNA to be packaged into viable phage particles. When unamplified phages were plated out at low density, we observed that there was a large variation in plaque size and that there were many very small plaques. As we have shown that the size of the phage plaque

can be affected by the DNA sequence cloned, amplification could potentially alter the relative frequencies of many DNA sequences. A major advantage of screening an unamplified library, is therefore, that poorly growing phages will be isolated which would be preferentially lost by amplification.

We isolated fewer 74K than 68K albumin recombinants probably for the following reasons. First, the background hybridisation to the nitrocellulose screening filters has to be extremely low to allow the detection of weak autoradiographic signals. Second, poorly growing phages are difficult to purify because they produce plaques containing up to 10^2 fewer pfu than normal. This means that, during the purification rounds, while large numbers of pfu's have to be plated out to be certain that the poorly growing phage is obtained, the plaques must be sufficiently spaced so that small plaques may be picked with a minimum of contamination. Thirdly, the titre of poorly growing recombinants decreased much faster than normal. This indicates that one reason why these phages grow slowly is because they produce unstable phage particles and necessitates that fresh high titre lysates are prepared to inoculate large scale cultures. It also suggests that DNA libraries should be plated for screening immediately after packaging to avoid the selective loss of unstable recombinants. A differential stability of phage particles due to the cloned DNA could also contribute to the unexpected frequencies of certain recombinants in amplified libraries which have been stored for a long period of time prior to screening. Finally, lysis of bacteria by slow growing phages in liquid culture takes 24-48h even when the ratio of phage to bacteria in the inoculum is dramatically increased. This slow growth presumably allows a strong selective pressure for phages containing deletions in their DNA which allow faster growth. Occasionally, double or multiple phage bands were observed after caesium chloride centrifugation of slow growing phages and restriction enzyme analysis of the X.laevis DNA isolated from the less dense phages showed it to be deleted. Although deletions seem to occur infrequently, these results suggest that recombinants should be stored as DNA which can be used to generate new phage particles by in vitro packaging when required.

ACKNOWLEDGEMENTS

We thank W. Wahli for the amplified X.laevis DNA library and advice on the construction and screening of the unamplified library, B. Hohn for providing bacterial strains, BHB 2688 and BHB 2690, G. Ryffel for discussion and E. Derungs for photography. F. M. thanks the Royal Society of Great Britain and B. W. the Roche Research Foundation for fellowships. This work was supported

by the Swiss National SCience Foundation.

*Present address: Inserm (U148), 60 rue de Navacelles, 34100 Montpellier, France.

+To whom reprint requests should be addressed.

REFERENCES

1. Wahli, W., Dawid, I.B., Wyler, T., Jaggi, R.B., Weber, R. and Ryffel, G.U. (1979) Cell 16, 535-549.
2. Widmer, H., Andreas, A-C., Weber, R. and Hosbach, H. (1982) Dev. Biol. (in press).
3. Bozzoni, I., Beccari, E., Luo, Z.X., Amaldi, F., Pierandrei-Amaldi, P. and Campioni, N. (1981) Nucl; Acid. Res. 9, 1069-1086.
4. Bisbee, C.A., Baker, M.A., Wilson, A.C., Hadji-Azami, I. and Fischberg, M. (1977) Science 195, 785-787.
5. Westley, B.R., Wyler, T., Ryffel, G.U. and Weber, R. (1981) Nucl. Acid. Res. 9, 3557-3574.
6. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) Cell 15, 687-701.
7. Wahli, W. and Dawid, I.B. (1980) Proc. Natl. Acad. Sci. USA. 77, 1437-1441.
8. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, I. (1977) J. Mol. Biol. 113, 237-251.
9. Ryffel, G.U. (1976) Eur. J. Biochem. 62, 417-423.
10. Davis, R.W., Thomas, M., Cameron, J., St. John, T.P., Scherer, S. and Padgett, R.A. (1980) in "Methods in Enzymolology", Colowick, S.P. and Kaplan, N.O. Eds., Vol 65, pp. 404-411, Academic Press, New York.
11. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
12. Wahli, W., Dawid, I.B., Ryffel, G.U. and Weber, R. (1981) Science 212; 298-304.
13. Sargent, E.F., Wu, J., Sala-Trepat, J.M., Wallace, R.B., Reyes, A.A.and Bonner, J. (1979) Proc. Natl. Acad. Sci. USA. 76, 3256-3260.
14. Kioussis, D., Eiferman, F., van de Rijn, P., Gorin, M.B., Ingram, R.S. and Tilghman, S.M. (1981) J. Biol. Chem. 256, 1960-1967.
15. Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.M., Slightom, J.L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980) Cell 21, 653-658.
16. Dolan, M., Sugarman, B.J., Dodgson, J.B. and Engel, J.D. (1981) Cell 24, 669-658.
17. Herrlich, P., Hynes, N.E., Ponta, H., Rahmsdorf, U., Kennedy, N. and Groner, B. (1981) Nucl. Acid. Res. 9, 4981-4995.
18. Fritsch, E.F., Lawn, R.M. and Maniatis, T. (1980) Cell 19, 959-972.