Volume 15 Number 11 1987 Nucleic Acids Research

Cosmid mapping of the human chorionic gonadotropin β subunit genes by field-inversion gel electrophoresis

Madge Y.Graham, Tetsuo Otani, Irving Boime, Maynard V.Olson[1], Georges F.Carle[1] and David D.Chaplin[2]

Departments of Pharmacology, [1]Genetics and [2]Medicine, and [2]the Howard Hughes Medical Institute, Washington University School of Medicine, St Louis, MO 63110, USA

Received February 27, 1987; Revised and Accepted May 12, 1987

ABSTRACT

A cosmid clone containing the entire hCGβ gene cluster has been isolated. The restriction map of this clone has been determined by an indirect-end-label FIGE (field inversion gel electrophoresis) method. Analysis of this cosmid clone shows that there are 6 hCGβ genes in human genomic DNA. A previously uncloned portion of the hCGβ cluster, termed the "gap" region, has been shown not to contain any sequences homologous to the hCGβ cDNA. The restriction mapping method employed in this study takes advantage of the superior resolution of FIGE for high molecular weight DNA fragments in the size range 15-50 kb. This method is broadly applicable and permits rapid and accurate restriction mapping for extended regions of genomic DNA that have been cloned into cosmid or lambda vectors.

INTRODUCTION

The placental hormone human chorionic gonadotropin (hCG) is a dimeric glycoprotein composed of two non-identical subunits, α and β. The alpha subunit, which is the product of a single-copy gene (1, 2), is common to hCG and the pituitary hormones lutropin (LH), follitropin, and thyrotropin (3). The beta subunits confer the functional specificity of the molecules. Analysis of the primary structures of the beta subunits of these hormones has shown hCG and LH (4, 5) to be the most similar. Recent studies have indicated that the genes encoding hCGβ and LHβ are clustered on a single EcoRI fragment of approximately 58 kb in length (6). LHβ is encoded by a single gene that lies at one end of the EcoRI fragment. In contrast, several hCGβ-related sequences, apparently representing both functional and non-functional genes, are distributed over most of the remainder of the fragment (6). The structural and functional organization of this gene cluster is of considerable interest since the presumed gene duplications that gave rise to it appear to have been followed by functionally significant divergences within coding and regulatory sequences (7). While LH and hCG have similar biologic activities, they are expressed in a tissue specific manner and their patterns of regulation are quite distinct. To understand the nature of the regulatory features of these

© IRL Press Limited, Oxford, England.

genes, detailed characterization of the molecular organization and structure of the cluster is necessary.

Previous studies of the LHβ/hCGβ gene cluster, based both on the use of cDNA probes to detect the coding regions in genomic restriction digests, and also on the analysis of genomic bacteriophage clones, have led to a model for the organization of the cluster that hypothesizes the presence of 6 coding regions for hCGβ (6). Final confirmation of this model has proven difficult, however, because a critical central region of the cluster is not present in any of the reported lambda clones; furthermore, the proposed 6-gene model depends on the assumption that some of the lambda clones from the region have artifactual structures (6-8).

In this paper, we describe a cosmid clone containing a 36.5 kb BamHI genomic insert that encodes all the genes for hCGβ. The cosmid clone was mapped by an indirect-end-label method that takes advantage of the high electrophoretic resolution of field-inversion gels (9) to overcome the problems that are normally encountered when mapping cosmid or lambda clones by end-label/partial-digestion techniques. By tuning the field inversions appropriately, high resolution can be achieved for fragments from 200 bp to 50 kb, allowing single-gel mapping of the entire cosmid with several restriction enzymes. Analysis of the cosmid clone indicates conclusively that the hCGβ gene cluster is defined by only 6 loci as suggested by prior Southern blot analysis (6).

MATERIALS AND METHODS

Preparation and Screening of Cosmid Library

Genomic DNA was prepared from the JAr choriocarcinoma cell line as previously described (1). It was digested to completion with BamHI and size fractionated using a Bull's Eye electrophoresis apparatus (Hoeffer) as previously described (10). DNA fragments of approximately 40 kb were pooled. The fractionated genomic DNA was ligated to cosmid vector pTCF (11) that had been linearized with BamHI and dephosphorylated using calf-intestinal phosphatase (Boehringer-Mannheim). Packaging, transduction, replica-plating, and screening were performed as previously described (10). The filters were hybridized with an hCGβ cDNA fragment (12) labeled with [^{32}P] by nick translation (13). A positive colony was isolated and cosmid DNA was prepared according to the method of Ish-Horowicz, *et al*. (14).

Partial Digestion, FIGE and Hybridization of Cosmid DNA

Cosmid DNA bearing hCGβ genomic sequences was first digested to completion with EcoRI, ethanol precipitated, and then subjected to partial

digestion with 0.2U/μg of HindIII, KpnI, XbaI or XhoI at 37°C for 5, 10, 20, or 30 min. Partial digests were monitored on a 0.6% agarose gel; for each enzyme, a 0.5μg sample of the digest with the broadest size distribution was selected for the field inversion gel. High molecular weight DNA standards from Bethesda Research Laboratories (BRL) and BstEII-digested lambda DNA were used as size markers for the field-inversion gel, which contained 1.0% agarose and was cast and run in 0.5 x TBE (1 x TBE = 89 mM Tris-borate, 89 mM boric acid and 2 mM EDTA). The gel was run with 14V/cm in the forward direction and 9.5V/cm in the reverse direction, alternatively applied each 0.3 sec. for 16h (9). The voltage differential was achieved using a single power supply and by placing an adjustable resistor in series with the gel when the reverse voltage was applied. The temperature of the gel was kept at 16°C by forced circulation of buffer through a heat exchanger (15). After electrophoresis, the DNA was transferred to a single sheet of nitrocellulose (16) which was then cut so that appropriate sets of lanes could be hybridized to the correct probe. One set of partial digests was hybridized to a [^{32}P]-labeled vector fragment adjacent to one of the EcoRI ends while an identical set was hybridized to a fragment adjacent to the other end (see below). Lambda markers were hybridized with [^{32}P]-labeled lambda DNA. Hybridizing bands were detected by autoradiography at -70°C using Kodak XAR film and a DuPont Lightning Plus screen.

Preparation of Cosmid Vector pTCF Fragments for Probes

pTCF DNA was digested with BamHI and HindIII to produce the 2.4 kb "right" end probe (Fig. 2) or BamHI and EcoRI to produce the 2.4 kb "left" end probe; the fragments were separated by electrophoresis in a 0.8% agarose gel in 0.5 x TBE, the desired gel fragments were excised from the gel, and the DNA fragments were electroeluted and purified on DE52 columns (17). They were labeled with [α-^{32}P]-dATP using a nick translation kit (BRL).

RESULTS

We prepared a cosmid library enriched for ~40 kb BamHI fragments derived from choriocarcinoma DNA (18). After transferring colonies to nitrocellulose filters and hybridizing with an hCGβ cDNA probe, we obtained a single clone that gave a positive signal. Cleavage patterns of this clone (termed βcos) with several enzymes are shown in Fig. 1A; the fragments that hybridize to an hCGβ coding-region probe are shown in Fig. 1B. These results confirm prior restriction fragment analysis of the hCGβ gene cluster which used genomic Southern blots and cloning of a portion of the region from recombinant bacteriophage libraries (6). Of particular significance, the cosmid clone

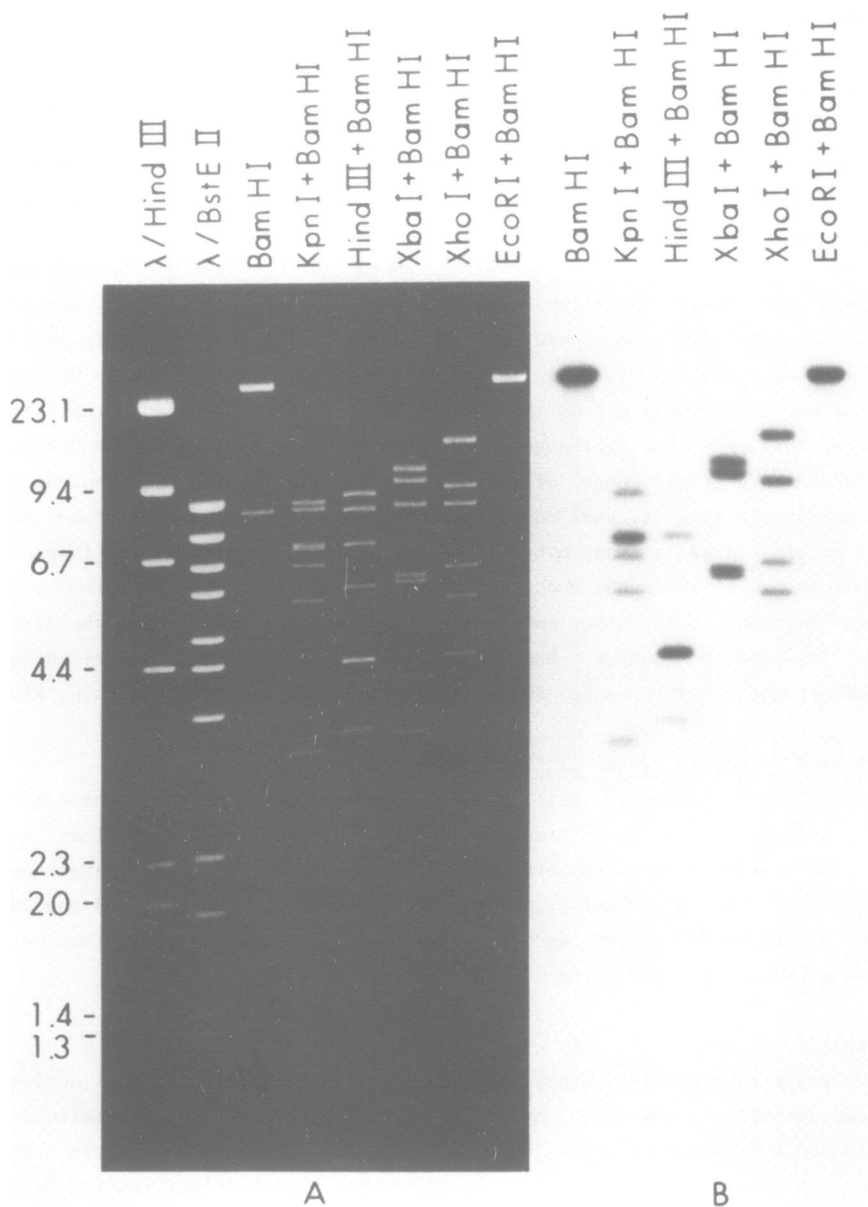


Fig. 1 Limit digest of a cosmid clone containing the entire hCGβ gene cluster (βcos) and hybridization analysis with hCGβ cDNA. The restriction endonucleases used are indicated above the lanes. Panel A, 0.6% agarose gel stained with ethidium bromide. Panel B, autoradiogram of the nitrocellulose filter transfer of Panel A which was hybridized with [^{32}P]-labeled pβ474 cDNA insert (12). The left margin shows marker fragment sizes in kilobase pairs.

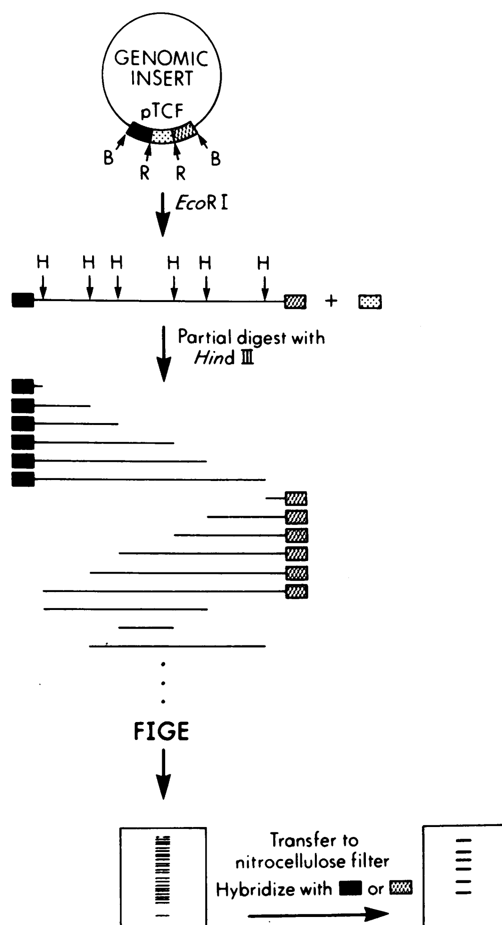
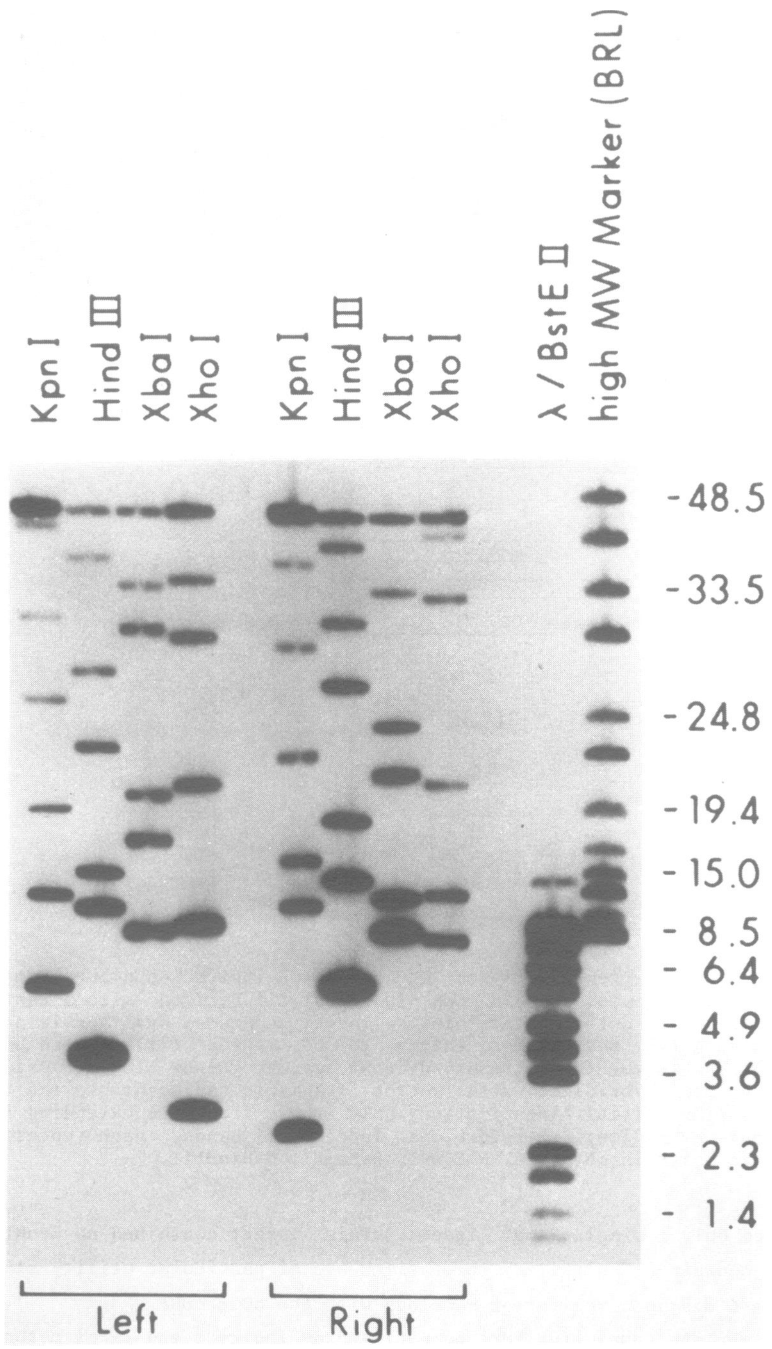


Fig. 2 Schematic representation of the FIGE indirect-end-label method for mapping cosmid clones. Cosmid DNA is first linearized with a restriction enzyme which does not digest within the insert sequences and then is partially digested with the restriction enzyme to be mapped (illustrated here with HindIII). The fragments are resolved on an agarose gel by FIGE, transferred to a membrane and hybridized with vector fragments adjacent to the two ends created during initial linearization. Only those fragments extending from the probed end hybridize; therefore, a ladder of bands, each representing a restriction site is obtained. R=EcoRI, B=BamHI, H=HindIII.

contained only a single BamHI insert. This insert contained no EcoRI sites, but did contain 6 KpnI sites; these sites are of particular utility since they separate 6 distinct regions of homology with the hCGβ cDNA.

To map the cosmid further, we employed an indirect-end-label method (Fig. 2). Since the cosmid insert contains no EcoRI sites, digestion with this


Kpn I
Hind III
Xba I
Xho I
Kpn I
Hind III
Xba I
Xho I
λ / BstE II
high MW Marker (BRL)
-48.5
-33.5
-24.8
-19.4
-15.0
- 8.5
- 6.4
- 4.9
- 3.6
- 2.3
- 1.4
Left
Right

enzyme generates a linear, insert-containing fragment with asymmetric vector sequences at the ends. Following limit EcoRI digestion, we performed partial digestions with KpnI, HindIII, XbaI, or XhoI. The fragments were separated on the basis of their lengths by FIGE, which was optimized to resolve fragments up to 50 kb, and then transferred to nitrocellulose. By hybridizing this filter with [^{32}P]-labeled probes derived from fragments adjacent to either the right or left EcoRI site, we obtained a ladder of hybridizing fragments (Fig. 3). The rungs of the ladder correspond to fragments extending from an EcoRI site in the vector to each restriction site in the insert; fragments not bounded by the EcoRI site from which the probe is derived do not hybridize and are not detected in the autoradiography.

A schematic of the experiment is shown in Fig. 2, and the hybridization results are shown in Fig. 3. Since fragments from both sides of the cloning site were used to hybridize duplicate filters, the analysis has an internal check; the sum of the distances separating a given restriction site from the two EcoRI sites should always be equal to the size of the intact EcoRI fragment. The indirect-end-label map derived this way had sufficient resolution to allow the assignment of the limit digest fragments (Fig. 1) to map positions in a straightforward way. The positions of the 5 KpnI/KpnI fragments and one KpnI/BamHI fragment that define the general location of the 6 hCGβ coding regions are shown in the middle of Fig. 4. We assigned the approximate location and orientation of each gene within these fragments according to our previous data with lambda clones and genomic Southern blots (6). Particularly informative were the HindIII sites; each hCGβ gene in the cosmid insert contains a HindIII site in the first intron.

The previously uncloned region in the center of the cluster lies entirely within a HindIII fragment that was subcloned from the cosmid into plasmid pβGAP (Fig. 4). Because a previously proposed 7-gene model for the hCGβ cluster (7, 8) requires an additional gene 5′-ward from gene 5, we analyzed pβGAP with an hCGβ coding-region probe. The plasmid was digested with MstII. This is a convenient enzyme since it generates the gap region free from the hybridizing exon sequences of genes 1 and 5 that flank the gap region. An MstII site is located about 180 bp upstream of the proposed cap site of gene 5

Fig. 3 Southern blot analysis of partially digested βcos DNA resolved by FIGE using end fragments of vector pTCF as probes. The restriction endonuclease used for partial digestion is shown above each lane. Two groups of lanes indicated as "left" and "right" are identical sets of partial digests hybridized with [^{32}P]-labeled vector fragments located immediately left or right of the BamHI cloning site. Lambda size markers were also transferred to the nitrocellulose filter and hybridized with [^{32}P]-labeled lambda DNA.

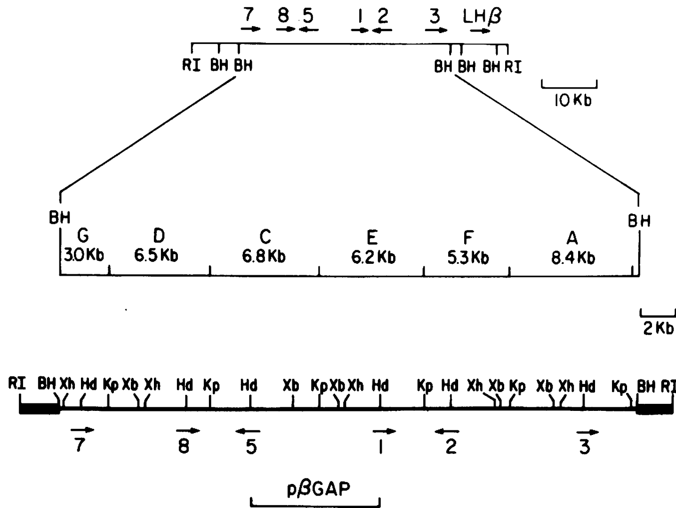


Fig. 4 Map of the hCGβ region. The top line indicates a composite map of the entire hCGβ/LHβ cluster showing the relative locations and orientations of the six hCGβ and the LHβ genes in the context of the BamHI and EcoRI sites. The bottom line indicates the map of the βcos cosmid as determined by the indirect-end-label FIGE method. Shaded EcoRI/BamHI sequences are derived from the cosmid vector. Assignment of fragments from a KpnI/BamHI limit digest (A to G in order of size) is shown in the middle line. The position of the HindIII fragment cloned as pβGAP, which spans a region most of which was previously uncloned, is indicated below the map.
BH=BamHI, Hd=HindIII, Kp=KpnI, RI=EcoRI, Xb=XbaI, Xh=XhoI.

(2). Another has been mapped to approximately the same location in the 5′-region of gene 1. Additionally, MstII cleaves two other sites within the gap region (data not shown). The only fragment in the digest that hybridizes to the hCGβ cDNA probe was the large fragment that contains vector sequences and the first exons of genes 1 and 5 (Fig. 5). These data indicate that there is no additional hCGβ gene in the gap region and confirm that there are only 6 hCGβ genes as suggested by Policastro, *et al.* (6).

DISCUSSION

Here, we report the use of FIGE to derive the restriction map of a cosmid clone containing the entire hCGβ gene family. Resolution of high molecular weight DNA was achieved to an extent which cannot be accomplished by conventional agarose gel electrophoresis, thereby overcoming a major limitation in previous applications of indirect-end-label mapping of cosmid and lambda clones (19-21). The cloned DNA was digested to completion with a restriction enzyme which cuts only in the vector and subsequently was digested

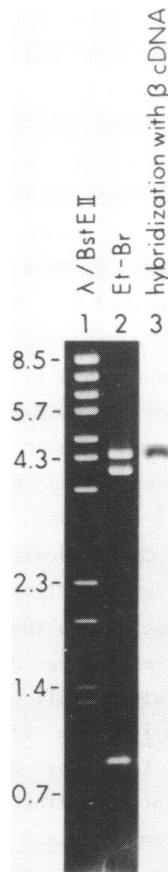


Fig. 5 Hybridization of an hCGβ cDNA to fragments generated by a limit double digest of pβGAP with MstII and XhoI. MstII cuts just 5′-ward of genes 1 and 5. Only the fragment containing exon 1 of both genes 1 and 5 and the vector pUC18 hybridized. No hybridization was observed in the fragments representing the sequence between genes 1 and 5.

partially with the enzymes used for restriction mapping. The resultant fragments were separated by FIGE, transferred to nitrocellulose, and hybridized to small fragments derived from each end of the vector. Only fragments extending from each EcoRI end of the vector are scored after hybridization. A ladder of hybridizing bands is obtained in which the distance of a particular restriction site from either EcoRI site can be determined. This method overcomes one of the most serious disadvantages of conventional electrophoretic techniques, the difficulty in ordering two or more closely adjacent restriction sites. With field inversion gels, the order of even quite

closely spaced restriction sites can be readily determined (see, for example, the redundant data documenting the left end...XbaI/XhoI... right end pair of sites 8.5 kb from the right end of the EcoRI fragment, lanes 3, 4, 7, and 8 in Fig. 3). Other forms of pulsed-field gel electrophoresis could undoubtedly also be used in this procedure. There is, however, little information in the literature on their performance in the 20-50 kb size range that is critical for lambda and cosmid mapping. In our hands, the highest mobility ratio that we have been able to achieve between a 48.5 kb and a 19.4 kb size marker by any transverse-field-alternation method has been 1.8, using a CHEF apparatus (22) with 1 sec switching. In contrast, the comparable ratio by field inversion using gels such as the one shown in Fig. 3 exceeds 3.

For the restriction mapping of the hCGβ-containing cosmid (βcos), we used digestion with EcoRI to generate a linear fragment containing the entire genomic insert bounded by vector-specified sequences. The pTCF cosmid vector also contains four other potential cleavage sites that could be used in an analogous way with other inserts: ClaI, NruI, and PvuI, and the bacteriophage lambda cos site (11). With the availability of these additional sites for linearization of the clone by digestion in the vector only, the method we have described is generally useful for most large genomic clones. In principle, the most generally applicable of the sites will be the lambda cos site, at which the cosmid can be linearized both in vitro (20) and in vivo (23). Since it is unlikely that cos sites will be found by chance in cosmid inserts, the cos site offers the possibility of pre-labeling the cleaved ends before running the mapping gel by annealing complementary kinased oligodeoxynucleotides to the 12-bp single-stranded overhang left by the lambda terminase reaction (19-21). Cross and Little (24) recently also reported the utility of field inversion in conjunction with the oligodeoxynucleotide pre-labeling protocol for cosmid mapping.

Previous studies by Policastro, et al. (6), have shown that all of the hCGβ gene sequences lie within one BamHI restriction fragment. This BamHI fragment was shown to be contained within a larger EcoRI fragment which also included the LHβ gene. We have used cosmid cloning to isolate the 36.5 kb BamHI fragment containing the hCGβ gene cluster. As predicted by the previous Southern blot analysis, this genomic insert contained no EcoRI sites, and it encoded 6 hCGβ-hybridizing sequences as well as the previously uncloned "gap" region. Analysis of the intact cosmid and of "gap"-specific subclones indicated that there were no sequences homologous to hCGβ in the "gap" region. This indicates that there are only 6 hCGβ genes in JAr DNA. Previous Southern blot analysis has shown no evidence of restriction fragment length

polymorphism between JAr and a panel of other human genomic DNA's (12).

In addition to providing an unambiguous restriction map of the hCGβ gene cluster, we anticipate that the βcos clone will be useful in characterizing the DNA sequences involved in tissue-specific and hormonally-regulated hCGβ gene expression. Previous efforts using hCGβ lambda bacteriophage clones to identify which hCGβ genes are transcriptionally active and to identify the hCGβ gene promoter(s) have been unsuccessful (Policastro, Ph.D. thesis). Gene transfer experiments utilizing the intact cosmid clone and various subclones may allow an assessment of the role of the "gap" region in the regulation of hCGβ gene expression; this region is of specific interest since it lies upstream to gene 5 which is probably a functionally expressed gene. In addition to gene 5, gene 3 and possibly gene 1 are also functional members of the cluster. These conclusions are based on comparison of the restriction maps of the cloned genes and of several hCGβ cDNA isolates as well as on expression of genes 3 and 5 in COS cells using SV-40-derived vectors (6, 25).

In summary, we have isolated a cosmid clone carrying all the hCGβ genes on a single 36.5 kb BamHI insert. FIGE has been applied to the mapping of restriction endonuclease cleavage sites within this clone. This method provides a rapid, reliable alternative to conventional mapping methods for analyzing extended genomic regions cloned using cosmid vectors. Detailed analysis of this cloned hCGβ-containing DNA has indicated that there are only 6 hCGβ genes and provides the material for a detailed analysis of the sequences required for transcription of functional hCGβ mRNA.

ACKNOWLEDGEMENT

This study was supported by grants from the Monsanto Company.

REFERENCES

1. Boothby, M., Ruddon, R.W., Anderson, C., McWilliams, D. and Boime, I. (1981). J. Biol. Chem. 256, 5121-5127.
2. Fiddes, J.C. and Goodman, H.M. (1981). J. Mole. Appl. Genetics 1, 3-18.
3. Pierce, J.G. and Parsons, T.F. (1981). Annu. Rev. Biochem. 50, 465-495.
4. Shome, B. and Parlow, A.F. (1973). J. Clin. Endocrinol. Metab. 36, 618-621.
5. Morgan, F.J., Birken, S. and Canfield, R.E. (1975). J. Biol. Chem. 250, 5247-5258.
6. Policastro, P., McQueen, S.D., Carle, G. and Boime, I. (1986). J. Biol. Chem. 261, 5907-5916.
7. Talmadge, K., Vamvakopoulos, N.C. and Fiddes, J.C. (1984). Nature 37-40.
8. Boorstein, W.R., Vamvakopoulos, N.C. and Fiddes, J.C. (1982). Nature 300, 419-422.
9. Carle, G.F., Frank, M. and Olson, M.V. (1986). Science 232, 65-68.
10. Chaplin, D.D., Woods, D.E., Whitehead, A.S., Goldberger, G., Colten,

H.R. and Seidman, J.G. (1983). Proc. Natl. Acad. Sci. U.S.A. 80, 6947-6951.
11. Grosveld, F.G., Lund, T., Murray, E.J., Mellor, A.L., Dahl, H.H. and Flavell, R.A. (1982). Nucleic Acids Res. 10, 6715-6732.
12. Policastro, P., Ovitt, C.E., Hoshina, M., Fukuoka, H., Boothby, M.R. and Boime, I. (1983). J. Biol. Chem. 258, 11492-11499.
13. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). J. Mol. Biol. 113, 237-251.
14. Ish-Horowicz, D. and Burke, J.F. (1981). Nucleic Acids Res. 9, 2989-2998.
15. Carle, G.F. and Olson, M.V. (1984). Nucleic Acids Res. 12, 5647-5664.
16. Southern, E.M. (1975). J. Mol. Biol. 98, 503-517.
17. Smith, H.O. (1980). Methods Enzymol. 65, 371-380.
18. Hussa, R.O., Story, M.T. and Pattillo, R.A. (1974). J. Clin. Endocrinol. Metab. 38, 338-340.
19. Rackwitz, H.R., Zehetner, G., Frischauf, A.M. and Lehrach, H. (1984). Gene 30, 195-200.
20. Rackwitz, H.R., Zehetner, G., Murialdo, H.D., Delius, H., Chai, J.H., Poustka, A., Frischauf, A. and Lehrach, H. (1985). Gene 40, 259-266.
21. Whittaker, P.A. and Southern, E.M. (1986). Gene 41, 129-134.
22. Chu, G., Vollrath, D. and Davis, R.W. (1986). Science 234, 1582-1585.
23. Little, P.F.R. and Cross, S.H. (1985). Proc. Natl. Acad. Sci. U.S.A. 82, 3159-3163.
24. Cross, S.H. and Little, P.F.R. (1986). Gene 49, 9-22.
25. Talmadge, K., Boorstein, W.R., Vamvakopoulos, N.C., Gething, M.-J. and Fiddes, J.C. (1984). Nucleic Acids Res. 12, 8415-8436.