found assembled with the petunia rbcL polypeptides, which are localized exclusively in the chloroplasts.

Although the catalytic site of Rubisco resides on rbcL, rbcS is also required for both carboxylase and oxygenase activities (*41*). The finding of hybrid holoenzymes in transformed petunia suggests the possibility of altering the enzymatic activity of this important enzyme by the introduction of novel genes coding for the rbcS polypeptide. Furthermore, the regulatory regions and transit sequence of the rbcS gene, when fused to a foreign gene, may be used to effect its light-induced expression and to target the gene product into chloroplasts. After the completion of this work, Murai *et al.* (*42*) demonstrated the expression of a bean phaseolin gene in sunflower cells after transfer via Ti plasmid vectors.

References and Notes

1. D. M. Coen, J. R. Bedbrook, L. Bogorad, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5487 (1977).
2. G. E. Blair and R. J. Ellis, *Biochim. Biophys. Acta* **319**, 223 (1973).
3. P. E. Highfield and R. J. Ellis, *Nature (London)* **271**, 420 (1978); N.-H. Chua and G. W. Schmidt, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 6110 (1978); A. R. Cashmore, M. K. Broadhurst, R. E. Gray, *ibid.* **75**, 655 (1978).
4. A. R. Grossman, S. Bartlett, N.-H. Chua, *Nature (London)* **285**, 625 (1980).
5. G. Coruzzi, R. Broglie, C. Edwards, N.-H. Chua, *EMBO J.*, in press.
6. G. Link, D. M. Coen, L. Bogorad, *Cell* **15**, 725 (1978).
7. S. M. Smith and R. J. Ellis, *J. Mol. Appl. Genet.* **1**, 127 (1982).
8. Y. Sasaki, T. Sakihama, T. Kamikubo, K. Shinozaki, *Eur. J. Biochem.* **133**, 617 (1983).
9. S. M. Smith and R. J. Ellis, *Nature (London)* **287**, 692 (1980).
10. R. Broglie, G. Bellemare, S. G. Bartlett, N.-H. Chua, A. R. Cashmore, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7304 (1981).
11. R. Broglie, G. Coruzzi, G. Lamppa, B. Keith, N.-H. Chua, *Biotechnology* **1**, 55 (1983).
12. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
13. G. Coruzzi, R. Broglie, A. R. Cashmore, N.-H. Chua, *J. Biol. Chem.* **258**, 1399 (1983).
14. S. L. Berry-Lowe, T. D. McKnight, D. M. Shah, R. B. Meagher, *J. Mol. Appl. Genet.* **1**, 483 (1982).
15. P. Dunsmuir, S. Smith, J. Bedbrook, *Nucleic Acid. Res.* **11**, 4177 (1983).
16. S. L. McKnight and R. Kingsbury, *Science* **217**, 316 (1982).
17. R. Fraley *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4803 (1983).
18. M.-D. Chilton *et al.*, *ibid.* **77**, 4060 (1980); N. Yadav *et al.*, *Nature (London)* **287**, 458 (1980); L. Willmitzer *et al.*, *ibid.*, p. 359.
19. L. Corvarrubias *et al.*, *Gene* **13**, 25 (1981).
20. F. Gosti-Testu, V. Norris, J. Brevet, *Plasmid* **10**, 96 (1983).
21. A. DePicker *et al.*, *J. Mol. Appl. Genet.* **1**, 561 (1982).
22. S. Rogers *et al.*, in preparation.
23. L. Marton *et al.*, *Nature (London)* **277**, 129 (1979).
24. G. Wullems, L. Molendijk, G. Ooms, R. Schilperoot, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 4344 (1981).
25. R. Horsch and G. Jones, *In Vitro* **16**, 103 (1980).
26. R. Fraley *et al.*, in preparation.
27. P. Thomas, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5201 (1980); G. G. Carmichael and G. R. McMaster, *Methods Enzymol.* **65**, 380 (1980).
28. T. Maniatis, A. Jeffrey, D. G. Kleid, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1184 (1975).
29. A. J. Berk and P. A. Sharp, *Cell* **12**, 721 (1977); R. F. Weaver and C. Weissman, *Nucleic Acids Res.* **7**, 1175 (1979).
30. P. J. Southern and P. Berg, *J. Mol. Appl. Genet.* **1**, 327 (1982).
31. L. Herrera-Estrella, A. Depicker, M. Van Montagu, J. Schell, *Nature (London)* **303**, 209 (1983).
32. D. A. Stetler and W. M. Laetsch, *Science* **149**, 1387 (1965).
33. A. Voller, D. E. Bidwell, A. Bartlett, *Bull. W. H. O.* **53**, 55 (1976).
34. G. Coruzzi, R. Broglie, G. Lamppa, N.-H. Chua, in *Structure and Function of Plant Genomes*, O. Ciferri and L. Dure, Eds. (Plenum, New York, 1983), pp. 47–59.
35. S. Werner and W. Machleidt, *Eur. J. Biochem.* **90**, 99 (1978).
36. P. Z. O'Farrell, H. M. Goodman, P. H. O'Farrell, *Cell* **12**, 1133 (1977).
37. S. M. Smith and R. J. Ellis, *Nature (London)* **278**, 662 (1979).
38. H. Roy, K. A. Costa, H. Adari, *Plant Sci. Lett.* **11**, 159 (1978).
39. M. W. Beven, R. B. Flavell, M.-D. Chilton, *Nature (London)* **304**, 184 (1983).
40. T. F. Gallagher and R. J. Ellis, *EMBO J.* **1**, 1493 (1982).
41. T. J. Andrews and B. Ballment, *J. Biol. Chem.* **258**, 7514 (1983).
42. N. Murai *et al.*, *Science* **222**, 476 (1983).
43. V. Glisen *et al.*, *Biochemistry* **13**, 2633 (1974).
44. M. L. Mishkind and G. W. Schmidt, *Plant Physiol.* **72**, 847 (1983).

45. We thank Dr. Carol Edwards for performing the ELISA experiments, Dr. Ben Dharm for preparing IgG–protein A–Sepharose beads, and Sara Adams, Nadera Ahmed, and Sarah Wagner for technical assistance. This work was supported in part by a Winston Foundation fellowship (to R.B.) and NIH postdoctoral fellowship GM07776 (to G.C.), and in part (for the work performed at the Rockefeller University) by grants GM-31500 and BRSG S17RR08065 from the National Institutes of Health and a grant from Monsanto Co.

26 January 1984; accepted 1 March 1984

RESEARCH ARTICLE

Expression Cloning of Human EGF Receptor Complementary DNA: Gene Amplification and Three Related Messenger RNA Products in A431 Cells

Chijen R. Lin, William S. Chen, Wiebe Kruiger
Lana S. Stolarsky, Wolfgang Weber, Ronald M. Evans
Inder M. Verma, Gordon N. Gill, Michael G. Rosenfeld

Epidermal growth factor (EGF), a peptide residue having 53 amino acids (*1*) derived from a 128-kD precursor predicted by recombinant DNA analysis (*2, 3*), effects both growth and induction of specific differentiated functions in target tissues (*4*). The actions of EGF are exerted through binding to a specific 170-kD plasma membrane glycoprotein receptor (*5–8*), which has a protein core of about 140 kD (*9–11*). The EGF receptor is a phosphoprotein (*12*) with intrinsic kinase activity specific for tyrosine residues (*13*); the receptor is autophosphorylated at a tyrosine residue (*12, 13*) and is regulated by phosphorylation at threonine and serine residues by the Ca^{2+}, phospholipid-dependent protein kinase (*14, 15*). A critical question has been to identify the mechanism or mechanisms by which binding of EGF to the plasma membrane receptor alters the phenotypic and growth properties of the cell.

EGF and thyrotropin-releasing hormone (TRH) produce a rapid (within minutes) stimulation of prolactin gene transcription (*16–18*), and specific 5′ flanking prolactin genomic sequences transfer EGF transcriptional regulation to other genes (*19*). These data suggest that rapid, gene-specific transcriptional regulation is likely to be one crucial determinant of EGF action. A series of genetic modifications of the EGF receptor would be useful for further dissecting molecular mechanisms by which the binding of EGF to its receptor regulates specific transcriptional effects.

C. R. Lin, W. S. Chen, L. Stolarsky, and M. G. Rosenfeld are members of the Eukaryotic Regulatory Biology Program, University of California, San Diego, School of Medicine, La Jolla 92093. W. Kruiger, R. M. Evans, and I. Verma are members of the Molecular Biology and Virology Laboratory, The Salk Institute, La Jolla, California 92037. W. Weber and G. N. Gill are members of the Division of Endocrinology, University of California, San Diego, School of Medicine, La Jolla 92093.

Abstract. *In order to further define the mechanisms by which polypeptide growth factors regulate gene transcription and cellular growth, expression cloning techniques were used to select human epidermal growth factor (EGF) receptor complementary DNA clones. The EGF 3′ coding domain shows striking homology to the transforming gene product of avian erythroblastosis virus (v-*erbB*). Over-expression of EGF receptors in A431 cell lines correlates with increased EGF receptor mRNA levels and amplification (up to 110 times) of the apparently singular EGF receptor gene. There appear to be three cytoplasmic polyadenylated RNA products of EGF receptor gene expression in A431 cells, one of which contains only 5′ (EGF binding domain) sequences and is postulated to encode the secreted EGF receptor-related protein.*

The overproduction of EGF receptors by a line of human epidermoid carcinoma cells (A431) (*20*) and evidence in these cells of chromosomal translocations involving the region of the EGF receptor genomic locus in chromosome 7 (*21*), suggested that characterization of the EGF receptor messenger RNA's (mRNA) and the encoding gene (or genes) might also give insight into specific transforming events. The observation that a 105-kD secreted protein exhibiting EGF binding (but no tyrosine kinase) properties is coproduced in A431 cells (*10*) raised the possibility that this protein might be encoded by mRNA distinct from EGF receptor mRNA.

We now report the cloning of DNA inserts complementary to portions of the coding region of EGF receptor mRNA. Sequence analysis of these complementary DNA (cDNA) clones revealed a marked homology of 3′ coding information with a series of structurally related retroviral encoded tyrosine kinases (*22–25*), and extraordinary sequence conservation compared to the avian erythroblastosis virus (AEV) transformation protein from v-*erb*B (*26–28*). The analysis of EGF receptor cDNA clones provides independent confirmation of the relation between the sequence of the EGF receptor protein and the v-*erb*B gene product reported by Downward *et al.* (*9*). In genomic DNA restriction map analysis, there appears to be a single EGF receptor gene, but three polyadenylated [poly(A)] mature transcripts (3.2, 6.3, and 10.6) are produced in A431 cells. While the 6.3- and 10.6-kb mRNA's contain both EGF binding and tyrosine kinase coding domains, the 3.2-kb mRNA contains only sequences hybridizing to 5′ (EGF binding domain) specific probes. A series of A431 clonal cell lines that overexpress the EGF receptor protein 2 to 100 times above that present in a normal fibroblast exhibit a coordinate 3- to 110-fold amplification of the EGF receptor gene and a commensurate increase in the three structurally related mRNA's.

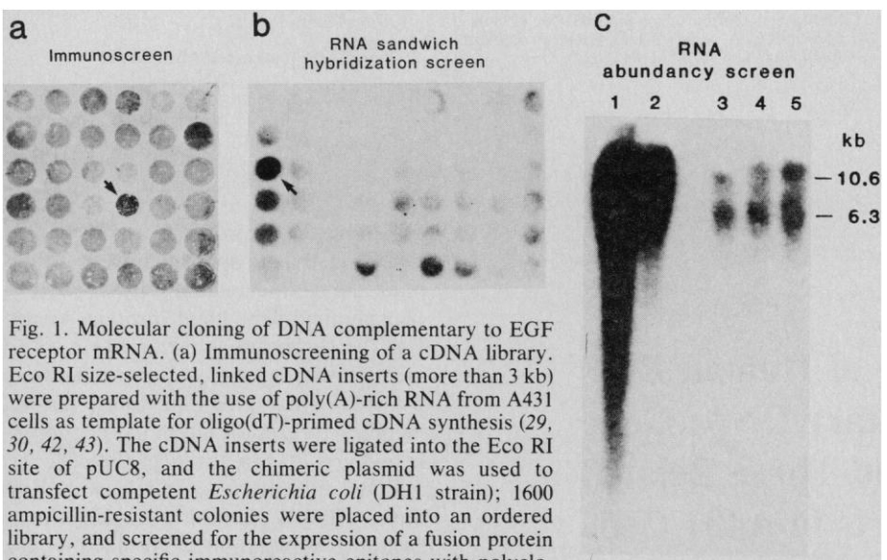


Fig. 1. Molecular cloning of DNA complementary to EGF receptor mRNA. (a) Immunoscreening of a cDNA library. Eco RI size-selected, linked cDNA inserts (more than 3 kb) were prepared with the use of poly(A)-rich RNA from A431 cells as template for oligo(dT)-primed cDNA synthesis (*29, 30, 42, 43*). The cDNA inserts were ligated into the Eco RI site of pUC8, and the chimeric plasmid was used to transfect competent *Escherichia coli* (DH1 strain); 1600 ampicillin-resistant colonies were placed into an ordered library, and screened for the expression of a fusion protein containing specific immunoreactive epitopes with polyclonal antiserum to purified EGF receptor protein from human A431 cells. Antiserums were affinity purified by adsorption to Sepharose-immobilized EGF receptor protein (*10*) with subsequent passage through a Sepharose column containing immobilized total soluble DH1 protein extract. The procedure was similar to that reported (*29*) except that horseradish peroxidase-coupled antiserum to rabbit immunoglobulin G (IgG) was used as second antibody. The darkest of four positive colonies is indicated by the arrow. Direct screening of the library with this probe revealed only three additional colonies. (b) Filter hybridization screening by means of an RNA sandwich hybridization assay. One positive cDNA insert was labeled with (3×10^8 cpm/μg) by nick-translation (*44*), hybridized to size-selected poly(A)-rich RNA from A431 cells ($R_0t = 0.7$) and the DNA · RNA hybrids were then used as probe in the colony filter hybridization assay (*45*). After a 12-hour hybridization and stringent washing (65°C, 0.1× SSC), 40 positive clones were visualized (for example, colony indicated by arrow). (c) RNA blot analysis screening. Poly(A)-selected RNA (3 μg) from A431 clone 29R cells (8×10^6 EGF receptors per cell, lane 1), A431 clone 1 cells (6×10^5 EGF receptors per cell, lane 2) human fibroblasts (about 1×10^5 EGF receptors per cell, lane 3), human epidermoid "L" cells (1.5×10^5 EGF receptors per cell, lane 4); human epidermoid "M" tumor cell line (3.5×10^5 EGF receptors per cell, lane 5) were separated by size under denaturing conditions, blotted, and hybridized to nick-translated cDNA inserts [in this figure the 760-bp insert of (pEGFR1), as previously described (*16, 46*)]. The radioautograph shows a 12-hour exposure; size standards were provided by the migration of a series of rat calcitonin mRNA species (*42*).

Cloning of cDNA Complementary to A431 EGF Receptor mRNA

The initial strategy used for the molecular cloning of DNA complementary to EGF receptor mRNA was based on the combined application of expression cloning techniques (*29, 30*), an RNA sandwich hybridization method devised for filter colony hybridization screening, and the availability of mRNA from a series of clonal epidermoid cancer cell lines (A431) which exhibit marked differences in EGF receptor content (*31, 32*). Poly(A) selected RNA prepared from one clonal A431 cell line (clone 29R, 8×10^6 receptors per cell) was used as template for avian myoblastosis virus (AMV) reverse transcriptase, and oligo deoxythymidylate was used as a primer (Fig. 1). Ampicillin-resistant transformants (1600) were placed in an ordered library and screened with polyclonal antibodies to EGF receptor protein purified to homogeneity from A431 cells. These antibodies appear to be directed at determinants in the EGF binding (amino-terminal) region of the receptor and specifically bind both the 170-kD holo EGF receptor protein and the 105-kD EGF receptor-related protein (ERRP), which lacks tyrosine kinase activity (*10*). Four colonies gave positive signals in this immunoscreening assay; an RNA sandwich technique permitted detection of additional potential positive colonies (Fig. 1). In this method, RNA is used as the probe to detect plasmids containing EGF receptor cDNA inserts, and the positive colonies

are visualized by hybridization of the RNA to clonal cDNA probes (for example, the immunopositive cDNA insert), such that noncontiguous cDNA clonal inserts can be detected. The criterion used to assess clonal inserts was their ability to differentially hybridize poly(A)-rich RNA from a series of cell lines producing EGF receptors at levels differing by 2 orders of magnitude. All inserts tested hybridized two large RNA species (6.3- and 10.6-kb), and the relative hybridization was in general accord with the content of EGF receptor protein. On the basis of the predictions by Downward *et al.* (*9*), we used an internal Bam HI fragment derived from v-*erb*B as probe to screen the library. A subset of ten chimeric plasmids (referred to as pEGFR1) contained an identically sized (760 bp) insert, presumed to represent an internal Eco RI fragment present in the double-stranded cDNA's, and hybridized to a v-*erb*B probe, suggesting that they represented cDNA's in the tyrosine kinase domain of the EGF receptor. The sizes of the two reactive RNA species are comparable to chicken embryo mRNA's hybridized by v-*erb*B probe (*33*). On the basis of these data the clones were further characterized as presumptive EGF receptor cDNA's.

The DNA sequence analysis (Fig. 2) demonstrates that pEGFR1 represents the EGF receptor-tyrosine kinase domain. The 762-bp fragment was found to have a single open reading frame. In accord with the data of Downward *et al.* (*9*), the encoded peptide has marked homology with the v-*erb*B transforming protein of AEV. In a region of 254 encoded amino acids, only six differ from v-*erb*B, a difference as likely to reflect species differences as any true divergence of v-*erb*B from primary EGF receptor sequence. This region spans the tyrosine residue thought to represent the autophosphorylation site in the kinase domain of EGF receptor (*12, 13*). The sequences predicted by the cDNA clonal analysis include sequences of peptide fragments reported by Downward *et al.* (*9*), and these independent methods of analysis exhibit precise sequence correspondence. In contrast to the extreme degree of peptide homology with v-*erb*B, the cDNA clone shows about 15 percent sequence divergence, although generally conservative, from the corresponding v-*erb*B coding domain. Additional cDNA's corresponding to both further 5′- and 3′-coding domains, have been identified. One clone (pEGFR3 1.8 kb insert) encompasses a portion of the EGF-binding domain and extends into the 5′ portion of the EGF receptor tyrosine kinase domain. There is no cross-hybridization between cDNA inserts representing particular 5′-coding and 3′-coding domains.

Analysis of genomic DNA from cells and tissues producing various quantities of EGF receptor reveals apparent structural identity of the 3′ portion of the gene in all tissues, as judged by the use of probes within the coding region. When the pEGFR1 insert was used as probe,

```
552
Lys Lys Ile Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Leu
AAA AAG ATC AAA GTG CTG GGC TCC GGT GCG TTC GGC ACG GTG TAT AAG GGA CTC
        G         T T         T   A   T   T       T A T               T
        Val                                         Ile
606
Trp Ile Pro Glu Gly Glu Lys Val Lys Ile Pro Val Ala Ile Lys Glu Leu Arg
TGG ATC CCA GAA GGT GAG AAA GTT AAA ATT CCC GTC GCT ATC AAG GAA TTA AGA
                  G   A   G               T   T       T   A       G
650
Glu Ala Thr Ser Pro Lys Ala Asn Lys Glu Ile Leu Asp Glu Ala Tyr Val Met
GAA GCA ACA TCT CCG AAA GCC AAC AAG GAA ATC CTC GAT GAA GCC TAC GTG ATG
  G   T       G   A                         A   T               T
704
Ala Ser Val Asp Asn Pro His Val Cys Arg Leu Leu Gly Ile Cys Leu Thr Ser
GCC AGC GTG GAC AAC CCC CAC GTG TGC CGC CTG CTT GGC ATC TGC CTC ACC TCC
  T   T   T       T   T   T             T     G   A               T
758
Thr Val Gln Leu Ile Thr Gln Leu Met Pro Phe Gly Cys Leu Leu Asp Tyr Val
ACC GTG CAA CTC ATC ACG CAG CTC ATG CCC TTC GGC TGC CTC CTG GAC TAT GTC
  T       G           C       T       T  AT                 T       C A
                                         Tyr                          Ile
812
Arg Glu His Lys Asp Asn Ile Gly Ser Gln Tyr Leu Leu Asn Trp Cys Val Gln
CGG GAA CAC AAA GAC AAT ATT GGC TCC CAG TAC CTG CTC AAC TGG TGT GTG CAG
  A   G       G       C                         T
866
Ile Ala Lys Gly Met Asn Tyr Leu Glu Asp Arg Arg Leu Val His Arg Asp Leu
ATC GCA AAG GGC ATG AAC TAC TTG GAG GAC CGT CGC TTG GTG CAC CGC GAC CTG
  T           A             C         A         C             T       T
                                      Glu
920
Ala Ala Arg Asn Val Leu Val Lys Thr Pro Gln His Val Lys Ile Thr Asp Phe
GCA GCC AGG AAC GTA CTG GTG AAA ACA CCG CAG CAT GTC AAG ATC ACA GAT TTT
  T               C   T   T   G   T   A   A       G   A           C
974                                               *
Gly Leu Ala Lys Leu Leu Gly Ala Glu Glu Lys Glu Tyr His Ala Glu Gly Gly
GGG CTG GCC AAA CTG CTG GGT GCG GAA GAG AAA GAA TAC CAT GCA GAA GGA GGC
          A   G       T   G   A   T       G   G   T   C       G
                                  Asp
1028
Lys Val Pro Ile Lys Trp Met Ala Leu Glu Ser Ile Leu His Arg Ile Tyr Thr
AAA GTG CCT ATC AAG TGG ATG GCA TTG GAA TCA ATT TTA CAC AGA ATC TAT ACC
  G   T       T   A                   G                   C     T       T
1082
His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Val Trp Glu Leu Met Thr Phe
CAC CAG AGT GAT GTC TGG AGC TAC GGG GTG ACC GTT TGG GAG TTG ATG ACC TTT
  T   A                   T   T   T       A                       A
1136
Gly Ser Lys Pro Tyr Asp Gly Ile Pro Ala Ser Glu Ile Ser Ser Ile Leu Glu
GGA TCC AAG CCA TAT GAC GGA ATC CCT GCC AGC GAG ATC TCC TCC ATC CTG GAG
  G           T           G       C   A   T   A               G   T
1190
Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Ile
AAA GGA GAA CGC CTC CCT CAG CCA CCC ATA TGT ACC ATC GAT GTC TAC ATG ATC
  G       G   T T G   C               T           T       G
1244
Met Val Lys Cys Trp Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Arg Glu Leu
ATG GTC AAG TGC TGG ATG ATA GAC GCA GAT AGT CGC CCA AAG TTC CGT GAG TTG
          A               T   T       C   C   T   C       T         C
```

Fig. 2. DNA sequence analysis of pEGFR1 cDNA insert. The 760-bp Eco RI excisable insert was mapped and found to contain one Sau 96 and two Bam HI sites; Eco RI, Bam HI, and Sau 96 digests were subjected to 3′ labeling with AMV reverse transcriptase full reactions and 5′ labeling with [γ-^{32}P]ATP and polynucleotide kinase (*47*). The sequence was confirmed by subcloning Alu and Eco RI fragments into an M13 vector (mp8) and the use of dideoxynucleotide DNA sequencing procedures (*48*). Nucleotide and amino acid divergences between the human pEGFR1 and the comparable v-*erb*B sequences (shown below the pEGFR1 sequence) are noted; the numbers used reflect the position in v-*erb*B sequence. The site of the autophosphorylated tyrosine residue is indicated by an asterisk.

three Eco RI and four Sac I reactive restriction products (Fig. 3) are seen. The identical Eco RI–generated bands are hybridized when an internal Bam HI restriction fragment of v-*erb*B is used as probe (data not shown). The number of reactive fragments hybridizing to a 760-bp probe is analogous to the findings in the chicken genome, where cellular-*erb*B (c-*erb*B) encompassed more than 20 kb (*33*). To evaluate the number of structurally distinct EGF receptor genes, a small Eco RI–Bam HI fragment (50 bp) from the 5′ end of pEGFR1 was used as probe. Because this probe hybridized only to the 6.9-kb Eco RI genomic fragment and hybridized to both the 6.3- and 10.6-kb mRNA's, it is likely that there is only a single EGF receptor gene. These data argue against the existence of a second gene encoding a protein with v-*erb*B homology but without an EGF binding domain; therefore, the EGF receptor gene would appear to represent c-*erb*B. The data do not exclude the existence of a second gene containing regions homologous only to 5′-coding EGF receptor domains.

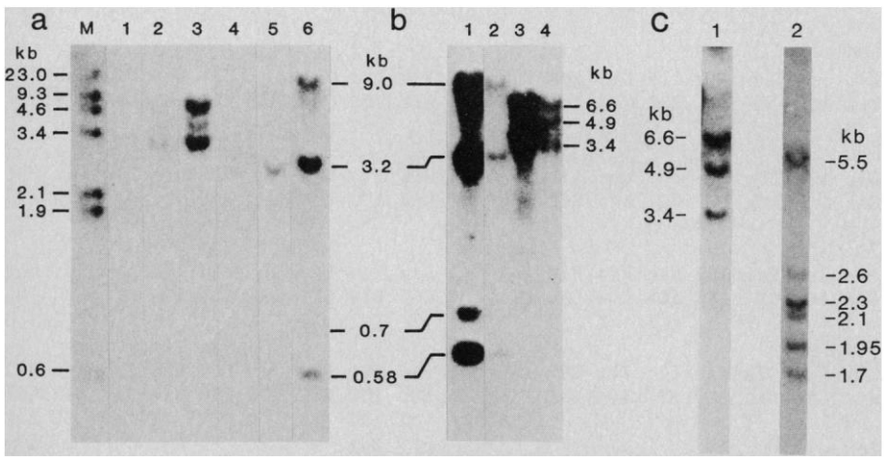
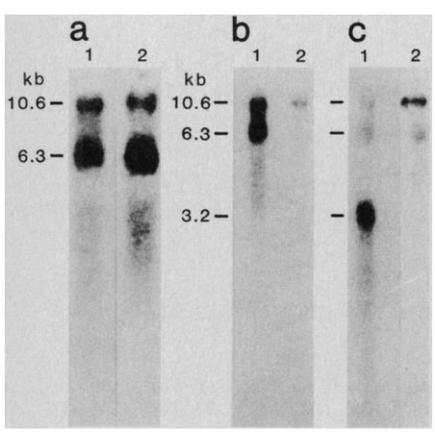


Fig. 3. Restriction map analysis of EGF receptor genomic DNA. Total nuclear DNA's from human placenta (panel a, lanes 1 and 4, and panel b, lanes 2 and 4), or A431 cell lines (clone 1: panel a, lanes 2 and 5, and panel b, lanes 1 and 3), and clone 29R (panel a, lanes 3 and 6) were digested with Eco RI (panel a, lanes 1 to 3, and panel b, lanes 3 and 4) or Sac I (panel a, lanes 4 to 6, and panel b, lanes 1 and 2). The fragments were separated by gel electrophoresis, transferred to nitrocellulose, and hybridized to nick-translated pEGFR1 insert (*49*). The size markers are provided by Hind III restriction fragments (lane M). Radioautographs were exposed for 3 hours (a) and 76 hours (b). In (c), Eco RI digested A431 genomic DNA was analyzed with the use of pEGFR1 (lane 1) or pEGFR3 (lane 2) as probe; sizes are based on the migration of lambda Hind III markers.



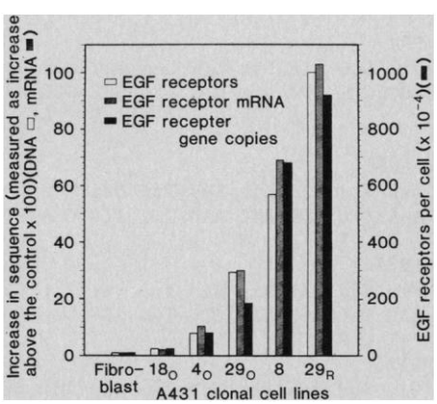


Fig. 4 (left). Characterization of EGF receptor cDNA-reactive mRNA's. Poly(A)-selected RNA (3 μg) from clone 29R A431 cells [cytoplasmic poly(A)-rich RNA, panel a, lane 1; total poly(A)-rich RNA, panel a, lane 2, and panels b and c, lane 1] or human fibroblasts (6 μg) (panels b and c, lane 2) was size-fractionated by denaturing, formaldehyde gel electrophoresis through 1.5 percent agarose gel transferred to nitrocellulose, and hybridized against nick-translated pEGFR1 (panels a and b) or pEGFR3 (panel c) probes. The radioautographs are exposed for 8 hours (a and b), 2 hours (panel c, lane 1), and 20 hours (panel c, lane 2), respectively. Size markers were provided by the migration of rat calcitonin precursors (*42*). The smear in panel a, lane 2, represents nonspecific background. Fig. 5 (right). EGF receptor gene amplification in A431 cell lines. Eco RI–digested genomic DNA and poly(A)-selected RNA, which were prepared from a series of A431 cell lines and human fibroblasts, were size-fractionated, blotted, and hybridized to nick-translated pEGFR1 insert as described in the legends to Figs. 3 and 4. The radioautographs were scanned, and the sum of the densitometric scan areas under the 6.4-, 4.3-, and 3.6-kb DNA restriction fragments and the 6.3-kb RNA were determined; fibroblast values were assigned 100 percent. The amount of EGF receptor protein was quantitated by affinity chromatographic purification and densitometric scanning of stained, electrophoretically separated EGF protein (*10*). The results are expressed as receptors per cell based on a protein molecular weight of 140 kD and using assumptions stated in (*10*).

Using clonal cDNA probes with different localizations in the EGF receptor coding domains, we analyzed the RNA products of EGF receptor gene expression by hybridization of denatured, size-fractionated poly(A)-selected RNA. Two cytoplasmic poly(A)-rich mRNA's, 6.3 and 10.6 kb in length, analogous to chicken c-*erb*B reactive RNA species (*33*), were detected with the clonal probe (pEGFR1) which was confined to sequences in the tyrosine kinase domain (Fig. 4). Both mRNA's were expressed in fibroblasts, a series of epidermoid tumors, and in A431 cells (Fig. 1). When a clonal probe which contained sequences encompassing both EGF binding and the 5′ portion of the tyrosine kinase domains (pEGFR3) was used as probe, a third poly(A)-rich RNA (3 to 3.2 kb in length) was visualized in A431 cells (Fig. 4). This RNA species, which is the most abundant, therefore contains sequences present in the EGF-binding domain, but not in the tyrosine kinase (v-*erb*B homologous) region. This cytoplasmic RNA might be expected to encode a protein with EGF binding, but not with tyrosine kinase activity, which is in precise accord with the biological properties of the secreted EGF receptor-related peptide (*10*). The 3.2-kb RNA is not detected in human fibroblast poly(A)-selected RNA. Analysis of appropriate cDNA clones suggests that the 3.2-kb mRNA arises consequent to use of a potential poly(A) site at an exon proximal to the sequences encoding the transmembrane-binding region of the EGF receptor. In addition, a series of five to six less reactive RNA species are consistently observed in A431 cells.

Amplification of EGF Receptor Gene in A431 cells

The EGF receptor cDNA clones permit quantification of both putative EGF receptor mRNA's and of the EGF receptor genes in cells overexpressing the EGF receptor. In epidermoid cell lines

overexpressing EGF receptor 1.5- to 3.5-fold, there is no alteration in gene copy number (data not shown); but EGF receptor mRNA does increase in proportion to increased receptor number (see Fig. 1C). In contrast, the A431 cells, which overexpress the EGF receptor 2- to 100-fold, exhibit a commensurate increase in the putative EGF receptor mRNA's and a 3- to 110-fold increase in the EGF receptor gene copy number (Fig. 5). Levels of the 3.2 kb RNA vary in proportion to the observed gene amplification characteristic of each cell line. Therefore, gene amplification could account for the increased EGF receptor production in these cells. Cytogenetic evidence suggests that translocations involving chromosome 7 occur in the region of the EGF receptor gene; however, neither homogeneous staining regions nor double minutes were observed (*21*). It is interesting that there are no apparent chromosomal rearrangements in some clonal cell lines exhibiting (eight- to tenfold) EGF receptor genomic amplification (Fig. 5, clone 40, and clone 1) (*34*) consistent with the known independence of genomic rearrangement and amplification events. Specific probes that represent 5′ and 3′ genomic flanking regions will be required to fully evaluate the question of gene rearrangement in A431 cells.

Implications of EGF Receptor mRNA Polymorphism

Several growth factor receptors, including those for EGF, insulin, and platelet-derived growth factor (PDGF) have intrinsic tyrosine kinase activity (*8, 35, 36*). The demonstration that known oncogenes with tyrosine kinase activity have considerable homology (*22–24*), and that they are structurally related to the catalytic subunit of cyclic adenosine monophosphate (cyclic AMP)–dependent protein kinase (*37*), suggested that the EGF receptor would be structurally related to this group of molecules. This prediction was confirmed by finding strong protein sequence homology between several cyanogen bromide cleavage fragments from human EGF receptor and v-*erb*B (*9*). The EGF receptor cDNA structural analysis presented above provides additional evidence for strong sequence homology with v-*erb*B as well as considerable homology to v-onc genes encoding tyrosine kinases. We present evidence for the existence of a single gene encoding the holo EGF receptor; therefore, the EGF receptor is postulated to represent c-*erb*B. In concert with the identification of homology between v-*sis* and PDGF (*38, 39*), these data support the notion that loss of regulation at any rate-limiting step in normal growth regulatory mechanisms could provide a growth advantage and initiate the process referred to as transformation.

From data on genomic DNA restriction map analysis, probes encompassing about 2.5 kb of coding region are represented in a 33-kb genomic region; therefore, the minimum size of the EGF receptor genome is larger than 40 kb. If there were a comparable organization across the entire gene, the transcription unit would be more than 100 kb, a transcription unit comparable to developmental genomic loci in *Drosophila*, and permitting potentially complex patterns of transcriptional and RNA processing regulation of EGF receptor gene expression. Consistent with this prediction, the EGF receptor gene appears to generate three discrete related transcripts in the A431 cell line. Two of the mature transcripts (6.3 and 10.6 kb) are present in all cells expressing EGF receptor and contain sequences present in both the EGF binding and tyrosine kinase domains of the receptor gene. Although the molecular basis for their diversity remains undefined, these RNA species appear to be expressed at variable levels in the cell types that we examined. Alternative transcription initiation site usage, RNA splicing events, or poly(A) site selection could account for their production. In this regard, it is interesting to note the tissue-specific pattern of alternative RNA processing events in certain genes of the neuroendocrine system (*40*). The third polyadenylated cytoplasmic RNA species (3.2 kb) contains only EGF binding domain-related sequences, and is considerably more prevalent than the two larger mRNA species. We propose that this 3.2-kb mRNA is likely to encode the recently isolated ERRP because ERRP contains EGF binding, but not tyrosine kinase, activity and is produced in equimolar quantities with the EGF receptor. Although this mRNA is not present in fibroblasts, it could be expressed in a tissue-specific or tumor-specific fashion. The 3.2-kb mRNA which apparently originates by alternative RNA processing events in EGF receptor gene expression could result from an allelic rearrangement of the receptor gene in A431 cells. On the basis of clonal cDNA analysis, we have identified a potential use of alternative RNA processing events in the case of EGF receptor gene expression involving a splice junction preceding by several amino acids the putative tyrosine autophosphorylation site (data not shown); therefore additional product diversity may be a feature of EGF receptor gene expression. A series of less reactive RNA species would represent processing intermediates, the homologous product of other genes, or functional mRNA products of EGF receptor gene expressions.

Overexpression of the EGF receptor protein in A431 cells reflects an increased EGF receptor mRNA content resultant from amplification of the EGF receptor gene. Such an amplification potentially provides a growth advantage to cells under conditions of limiting EGF. Although EGF inhibits growth of parental A431 cells (*41*), in many clones and under most conditions, it potentiates cell growth (*32*). Doubling times of clonal A431 cell variants maintained in 0.5 percent calf serum are inversely related to the concentration of EGF receptor protein (data not shown). Although translocation of chromosome 7 in the region postulated to include the EGF receptor gene has been reported (*21*), two A431 clonal lines (4_O and 1), in which no chromosomal translocations have been detected (*34*), still show at least a threefold amplification of the EGF receptor gene amplification.

The use of expression cloning and an RNA sandwich screening approach has made available EGF receptor cDNA which permits experiments designed to evaluate structure-function relations between receptor-ligand interaction, specific gene transcription, and growth regulation.

References and Notes

1. S. Cohen, *J. Biol. Chem.* **237**, 1555 (1962).
2. A. Gray, T. J. Dull, A. Ullrich, *Nature (London)* **303**, 722 (1983).
3. J. Scott, M. Urdea, M. Quiroga, R. Sanchez-Pescador, N. Fong, M. Selby, W. J. Rutter, G. I. Bell, *Science* **221**, 236 (1983).
4. G. Carpenter and S. Cohen, *Annu. Rev. Biochem.* **48**, 193 (1979).
5. I. Covelli, R. Rossi, R. Mozz, L. Fratti, *Eur. J. Biochem.* **27**, 225 (1972).
6. M. D. Hollenberg and P. Cuatrecasas, *Proc. Natl. Acad. Sci. U.S.A.* **70**, 2964 (1973).
7. G. Carpenter, K. J. Lembach, M. Morrison, S. Cohen, *J. Biol. Chem.* **250**, 4297 (1975).
8. S. Cohen, H. Ushiro, C. Stoscheck, M. Chinkers, *ibid.* **257**, 1523 (1982).
9. J. Downward *et al.*, *Nature (London)* **307**, 521 (1984).
10. W. Weber, G. N. Gill, J. Spiess, *Science* **224**, 294 (1984).
11. G. Carpenter, *Mol. Cell. Endocrinol.* **31**, 1 (1983).
12. T. Hunter and J. A. Cooper, *Cell* **24**, 741 (1981).
13. H. Ushiro and S. Cohen, *J. Biol. Chem.* **255**, 8363 (1980).
14. C. Cochet, G. N. Gill, J. Meisenhelder, J. A. Cooper, T. Hunter, *ibid.* **259**, 2553 (1984).
15. S. Iwashita and C. F. Fox, *ibid.*, p. 2559.
16. G. H. Murdoch, E. Potter, A. K. Nicholaisen, R. M. Evans, M. G. Rosenfeld, *Nature (London)* **300**, 192 (1982).
17. G. H. Murdoch, M. G. Rosenfeld, R. M. Evans, *Science* **218**, 1315 (1982).
18. G. H. Murdoch, R. Franco, R. M. Evans, M. G. Rosenfeld, *J. Biol. Chem.* **258**, 15329 (1983).
19. S. Supowit, E. Potter, R. M. Evans, M. G.

Rosenfeld, *Proc. Natl. Acad. Sci. U.S.A.*, in press.
20. R. N. Fabricant, J. E. DeLarco, G. J. Todaro, *ibid.* **74**, 565 (1977).
21. N. Shimizu, I. Kondo, S. Gamou, M. A. Behzadian, Y. Shimizu, *Somatic Cell Mol. Genet.* **10**, 145 (1984).
22. M. Shibuya and H. Hanafusa, *Cell* **30**, 787 (1982).
23. A. Hampe, I. Laprevotte, F. Galibert, L. A. Fedele, C. J. Scherr, *ibid.*, p. 775.
24. N. Kitamura, A. Kitamura, K. Toyoshima, Y. Hirayama, M. Yoshida, *Nature (London)* **297**, 205 (1982).
25. R. Muller and I. M. Verma, *Curr. Top. Immunol. Microbiol.*, in press.
26. Y. Yamamoto, T. Nishida, N. Miyajima, S. Kawai, K. Ooi, K. Toyoshima, *Cell* **35**, 71 (1983).
27. M. L. Privalsky, T. Ralston, J. M. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 704 (1984).
28. L. Frykberg, S. Palmieri, H. Beug, T. Graf, M. J. Hayman, B. Vennstrom, *Cell* **32**, 227 (1983).
29. D. M. Helfman, J. R. Faramisco, J. C. Fiddes, G. P. Thomas, S. H. Hughes, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 31 (1983).
30. K. E. Mayo, W. Vale, J. Rivier, M. G. Rosenfeld, R. M. Evans, *Nature (London)* **306**, 86 (1983).
31. J. E. Buss, J. E. Kudlow, C. S. Lazar, G. N. Gill, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2574 (1982).
32. T. Kawamoto, J. Mendelsohn, A. Le, G. H. Sato, C. S. Lazar, G. N. Gill, *J. Biol. Chem.*, in press.
33. B. Vennstrom and J. M. Bishop, *Cell* **28**, 135 (1982).
34. N. Shimizu, S. Gamou, D. M. Thompson, W. Weber, H. Masui, G. N. Gill, unpublished data.
35. M. Kasuga, Y. Fujita-Yamaguchi, D. L. Blithe, C. R. Kahn, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2137 (1983).
36. G. Ek and C.-H. Heldin, *J. Biol. Chem.* **257**, 10486 (1982).
37. W. C. Barker and M. O. Dayhoff, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2836 (1982).
38. R. F. Doolittle, *Science* **221**, 275 (1983).
39. M. D. Waterfield *et al.*, *Nature (London)* **304**, 35 (1983).
40. M. G. Rosenfeld *et al.*, *ibid.*, p. 129.
41. G. N. Gill and C. S. Lazar, *ibid.* **293**, 305 (1981).
42. S. G. Amara, V. Jonas, M. G. Rosenfeld, E. S. Ong, R. M. Evans, *ibid.* **298**, 240 (1982).
43. R. Bolivar, R. L. Rodriguez, P. J. Greene, M. C. Betlack, H. L. Hayuneker, H. W. Boyer, *Gene* **2**, 95 (1977).
44. P. W. J. Rigby, M. Dieckmann, C. Rhodes, P. Berg, *J. Mol. Biol.* **113**, 237 (1977).
45. M. Grunstein and D. S. Hogness, *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3961 (1975).
46. G. M. Wahl, R. A. Padgett, G. R. Stark, *J. Biol. Chem.* **254**, 8679 (1979).
47. A. M. Maxam and W. Gilbert, *Methods Enzymol.* **65**, 499 (1980).
48. F. Sanger, S. Nicklen, A. Coulsen, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5467 (1977).
49. E. M. Southern, *J. Mol. Biol.* **98**, 503 (1975).
50. We thank Cheri Lazar, Vivian Jonas, and Debbie Finley for their research contributions and suggestions and Dr. James Wedner for the preparation of EGF receptor antiserum. Supported by NIH grants, a Council for Tobacco Research grant, an American Cancer Society grant, a grant from Deutsche Forschungsgemeinschaft (W.W.), and a fellowship of the Dutch Queen Wilhelmina Fund (W.K.).

5 April 1984; accepted 27 April 1984

AAAS–Newcomb Cleveland Prize
To Be Awarded for an Article or a Report Published in *Science*

The AAAS–Newcomb Cleveland Prize is awarded annually to the author of an outstanding paper published in *Science*. The 1984 competition starts with the 6 January 1984 issue of *Science* and ends with the issue of 21 December 1984. The value of the prize is $5000; the winner also receives a bronze medal.

Reports and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration for the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the year, readers are invited to nominate papers appearing in the Reports or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS–Newcomb Cleveland Prize, AAAS, 1515 Massachusetts Avenue, NW, Washington, D.C. 20005. Final selection will rest with a panel of distinguished scientists appointed by the Board of Directors.

The award will be presented at a session of the AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.