

geometry of imbricate slices within the duplex versus >234 km displacement of the roof. Consequently Mount Isa Group sediments from either side of the roof thrust were originally deposited hundreds of kilometres apart. However, since the tectonic structural controls on sedimentation were apparently also north-south oriented the effect is not quite as critical as if they had been east-west oriented.

The distribution of the Fiery Creek Volcanics, for example, is considerably simplified by the thrust interpretation. They occur on the roof thrust and those imbricates to the north end of the duplex, which explains their absence in the imbricate slices farther to the south yet their presence to the east and west.

The significance of this interpretation to exploration for Mount Isa style Pb-Zn and copper orebodies is fundamental, especially considering the origin of the copper^{14,24,25} at Mount Isa relative to the structural history. This information cannot be discussed yet as it is waiting for company approval.

The geology of this region, if the thrust interpretation is correct, will enable us to develop an understanding of thrust

belt geometry and processes that is unparalleled. Here we have a large scale thrust belt in three dimensions due to the folding effects of the second deformation. At least 15 km of cross-section through this crust belt is exposed in the core of the anticlinal structure and this has already required previously undescribed thrust geometries to be predicted, modelled and applied. Ready explanations for cross-sectional relationships previously inexplicable in terms of accepted models of duplex development¹⁸⁻²⁰ have also been developed. The structural relationships also indicate that duplexes can form by thrusting on different stratigraphical levels above and below the package of horses. These relationships and the results of further work on this area are of considerable interest, especially in the light of recent COCORP^{27,28} revelations on the significance of deep levels of thrusting in the tectonic development of various regions in the USA such as the Appalachians.

I thank Andy Duncan and Bill Perkins for helpful discussions and Dave Elliott (now deceased) for rekindling my interest in thrust belts.

Received 8 April; accepted 16 June 1983.

1. Smith, W. D. *Spec. Publ. Geol. Soc. Aust.* **3**, 225-235 (1969).
2. Dunnet, D. *Phil. Trans. R. Soc. A* **283**, 333-344 (1976).
3. Derrick, G. M. *BMR J. Aust. Geol. Geophys.* **7**, 81-92 (1982).
4. Winsor, C. N. thesis, James Cook Univ. (1982).
5. Winsor, C. N. *Tectonophysics* **92**, 195-210 (1983).
6. Winsor, C. N. *J. geol. Soc. Aust.* (submitted).
7. Hutton, I. et al. *BMR Aust. Prelim. Map Sheet* 6758 (1982).
8. Derrick, G. M. & Wilson, I. H. *BMR Aust. Rec.* 1982/6 (1982).
9. Derrick, G. M., Wilson, I. H., Hill, R. M., Glikson, A. Y., Mitchell, J. E. *BMR Aust. Bull.* 193 (1977).
10. Wilson, I. H. et al. *BMR Aust. Rec.* 1977/4 (1977).
11. Wilson, I. H., Hill, R. M., Noon, T. A., Duff, B. A. & Derrick, G. M. *BMR Aust. Rec.* 1979/24 (1979).
12. Wilson, I. H. et al. *BMR Aust. Prelim. Map Sheet* 6859 (1981).
13. Derrick, G. M. et al. *BMR Aust. Prelim. Map Sheet* 6759/6760 (1980).
14. Swager, C. P. thesis, James Cook Univ. (1983).
15. Winsor, C. N. *J. struct. Geol.* (submitted).
16. Bell, T. H. & Duncan, A. C. *Tectonophysics* **47**, T1-T5 (1978).
17. Willis, B. *Bull. geol. Soc. Am.* **13**, 305-352 (1902).
18. Boyer, S. E. & Elliott, D. *Am. Ass. petrol. Geol.* **66**, 1196-1230 (1982).
19. Elliott, D. & Johnson, M. R. W. *Trans. R. Soc. Edinb.* **71**, 69-96 (1980).
20. Butler, R. W. H. *J. struct. Geol.* **4**, 239-245 (1982).
21. Dahlstrom, C. D. A. *Bull. Can. petrol. Geol.* **18**, 332-406 (1970).
22. Bally, A. W. *Phil. Trans. R. Soc. A* **305**, 325-338 (1982).
23. Bally, A. W. *Oceanol. Acta, Proc. 26th int. Congr.* 87-101 (1981).
24. Perkins, W. G. *Mt Isa Mines Ltd. Int. Rep.* 120/2 (1979).
25. Perkins, W. G. *Econ. Geol.* (submitted).
26. Page, R. W. *J. geol. Soc. Aust.* **25**, 141-164 (1978).
27. Cook, F. A. et al. *Geology* **7**, 563-567 (1979).
28. Cook, F. A. & Oliver, J. E. *Am. J. Sci.* **281**, 993-1008 (1981).

Structure of the Ki-ras gene of the human lung carcinoma cell line Calu-1

Kenji Shimizu*, Daniel Birnbaum, Mary Ann Ruley, Ottavio Fasano*, Yolande Suard, Louise Edlund, Elizabeth Taparowsky, Mitchell Goldfarb & Michael Wigler

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

The homologue of the viral Kirsten ras (v-Ki-ras) gene found in the human lung carcinoma cell line, Calu-1, has an intron-exon structure similar to that of the human homologue of the viral Harvey ras (v-Ha-ras) gene. A second, potential fourth coding exon is present in the human Ki-ras gene and similar sequences are found in the Kirsten murine sarcoma virus. Cysteine is encoded at the twelfth amino acid position, suggesting that the Calu-1 Ki-ras gene has undergone a mutational activation at the same position as the human Ha-ras gene of the bladder carcinoma cell line, T24. A comparison of their predicted amino acid sequences suggests that ras proteins have a 'constant' region and a 'variable' region. Here we propose a common modular structure for ras gene products in which the variable region forms a physiologically important combining site.

THE *ras* genes were first characterized as the transforming genes of the Harvey and Kirsten murine sarcoma viruses (v-Ha-ras and v-Ki-ras, respectively)¹. These viral *ras* genes share considerable nucleotide homology and encode structurally

related 21,000-molecular weight proteins (p21)¹. Each arose by transduction of a distinct cellular rat gene, and is highly conserved in evolution¹. Use of DNA-mediated gene transfer has shown that DNAs from many human tumours and human tumour cell lines contain *ras* genes capable of tumorigenic and morphological transformation of NIH 3T3 murine fibroblasts²⁻⁵. The transforming gene of T24 bladder carcinoma cells is the human homologue of the viral Ha-ras²⁻⁴. A transforming gene commonly found in human carcinomas of lung, colon and

* Present address: Department of Biology, Faculty of Science, Kyushu University, Fukuoka 812, Japan (K.S.). Permanent address: I Istituto di Chimico Biologica, II Facoltà di Medicina, Università di Napoli, Italy (O.F.).

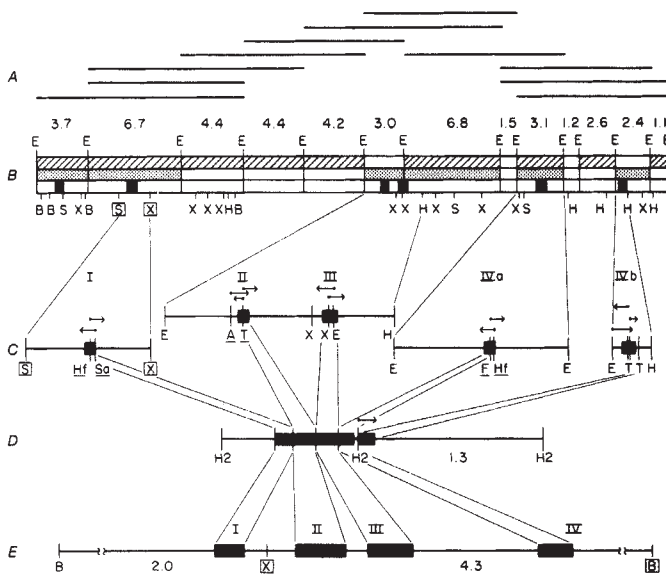


Fig. 1 *A*, Representative overlapping phage inserts from Charon 4A libraries containing portions of the human *Ki-ras* gene cloned from NIH 3T3 cells transformed with Calu-1 DNA. The restriction map of the *Ki-ras* gene is shown in *B*. Hatched regions indicate *EcoRI* fragments having homology to human repeat sequences as determined by strong hybridization with ^{32}P -labelled nick-translated human DNA²¹. Stippled regions indicate *EcoRI* fragments having homology to the pKBE-2 clone of *Ki-MuSV*⁷. Solid regions indicate more precisely the region of viral *Ki-ras* homology. *C*, An expansion of the regions of viral *Ki-ras* homology showing the strategy used in nucleotide sequence determination. Arrows indicate regions sequenced by the method of Maxam and Gilbert²² after 3' OH labelling²³. *D*, Map of the viral *Ki-ras* gene. The regions of homology between the human and viral genes are indicated, as is the strategy used to sequence the alternate IVb coding exon of the viral *Ki-ras*. *E* Shows the structure of the human *Ha-ras* gene, as determined previously^{11,12}. Solid boxes indicate coding regions. Regions of homology between the human *Ha-ras* and viral *Ki-ras* are indicated. Numbers indicate the distance in kbp. Restriction endonuclease sites are: E, *EcoRI*; B, *BamHI*; S, *SmaI*; X, *XbaI*; H, *HindIII*; Hf, *HinfI*; Sa, *Sau3A*; A, *AhaIII*; T, *TaqI*; F, *FokI*; H2, *HincII*. Restriction sites that are underlined are not unique for the restriction fragment shown. Restriction sites in boxes frame DNA fragments that were used in the chimaeric gene construction described in the text.

other tissues is the homology of the viral *Ki-ras*^{2,5}. The transforming gene of the SK-N-SH neuroblastoma cell line is called *N-ras*; it encodes a protein structurally related to the other *ras* gene products⁵. These transforming genes presumably arise by a somatic mutational 'activation' of the respective cellular *ras* genes⁶⁻⁸.

Cloning the *Ki-ras* gene

We have cloned the transforming *Ki-ras* gene of Calu-1, a human lung carcinoma cell line⁹. Primary transformants of NIH 3T3 cells were generated by gene transfer with Calu-1 DNA. DNA from primary transformants was then used to generate secondary and, in turn, tertiary, NIH 3T3 transformants. Phage libraries were constructed in Charon 4A from secondary and tertiary transformant DNAs which had been partially cleaved with the restriction endonuclease *EcoRI*. The first *Ki-ras* phage clone was obtained by screening these libraries with the human repeat probe, BLUR8, as briefly described elsewhere⁵. This clone, AL2-34, was cleaved with restriction endonucleases, and restriction fragments of phage insert were used for further screening of phage libraries constructed from DNA of NIH 3T3

transformants. This process was reiterated. In this manner we obtained 34 phage with overlapping restriction maps spanning 45 kilobase pairs (kbp). This entire region was mapped for *XbaI*, *SmaI*, *BamHI*, *HindIII* and *EcoRI* restriction endonuclease cleavage sites, and resulting DNA fragments were blotted and screened for homology to human repeat sequence and viral *Ki-ras* sequences. The results are shown in Fig. 1. We obtained all the *EcoRI* fragments containing human repeat sequences which are reproducibly associated with this gene on gene transfer (in kbp): 6.8, 6.7, 4.4, 4.2, 3.7, 3.1, 2.6 and 1.1 (ref. 9). Similarly, we obtained all *EcoRI* fragments having homology to a plasmid clone of viral *Ki-ras*, pKBE-2 (ref. 1), which are reproducibly associated with this gene (in kbp): 6.8, 6.7, 3.1, 3.0 and 2.4 (ref. 5). One novel *EcoRI* fragment of 3.7 kbp, which had not been previously appreciated from genomic blots with viral *Ki-ras* probes, had homology to sequences 5' to viral *Ki-ras* coding sequences (data not shown). Finally, the 2.4-kbp *EcoRI* fragment hybridized to a region of the Kirsten murine sarcoma virus (*Ki-MuSV*) 3' to the canonical *ras* coding sequences. We conclude that we have cloned an intact human *Ki-ras* gene.

Ki-ras coding regions

We prepared a finer restriction endonuclease map of all regions having homology to the viral *Ki-ras* gene. These regions were sequenced as described in Fig. 1 legend. Comparison of the predicted amino acid sequence of the *Ki-ras* with that of the published viral *Ki-ras*¹⁰ allowed an unambiguous assignment of the coding exons. Four coding exons (labelled I, II, III and IVa in Fig. 1) together would comprise an 189-amino acid protein product differing from the viral *Ki-ras* p21 at only seven positions (see Fig. 2). Consensus splice sequences were found at each intron-exon boundary (data not shown). The placement of these boundaries relative to the encoded protein corresponded exactly to the intron-exon boundaries of the human *Ha-ras* gene^{11,12} (see Fig. 2), consistent with the notion that *Ha-ras* and *Ki-ras* both evolved from a common, similarly spliced ancestral gene.

In addition to the first four coding exons, we found a potential coding exon (labelled IVb in Fig. 1) having homology to a region 3' to the termination codon of the viral *Ki-ras* gene. Both the human and viral regions were sequenced. The *Ki-ras* IVb sequence is preceded by a consensus splice sequence, and consists of an open reading frame which would translate into 38 amino acids, one less than would be found in *Ki-ras* IVa. The *Ki-ras* IVa- and IVb-encoded amino acid sequences are also quite similar, encoding a related 14-amino acid sequence at the 5' end of the exon and a related 5-amino acid sequence at the 3' end (see Fig. 3). The intervening amino acids (20 for *Ki-ras* IVa and 19 for *Ki-ras* IVb) are highly divergent. These two exons (IVa and IVb) have diverged from each other in nearly the same manner as the fourth exons of *N-ras*, *Ha-ras* and *Ki-ras* have each diverged from each other (see Fig. 3). Analysis of cDNA clones from human *Ki-ras* transformants indicates that both the IVa and IVb exonic sequences are expressed (M.G., unpublished results). Precedent exists for multiple gene products due to differential splicing of transcripts from the same genetic locus in vertebrates¹³⁻¹⁵.

Examination of the viral *Ki-ras* IVb region indicates that sequences similar to the human *Ki-ras* IVb must exist in the rat genome. Viral *Ki-ras* IVb sequences begin 5 base pairs (bp) 3' to the *Ki-ras* termination codon. The human and viral IVb sequences encode a similar amino acid sequence up to and including a poly-lysine stretch (Fig. 3). After this, there appears to be complete divergence of amino acid sequence. However, this part of the viral *Ki-ras* IVb can also be aligned with the human *Ki-ras* IVb by a 5-bp deletion within the oligo(dA) stretch (Fig. 3), yielding again a strikingly similar amino acid sequence. We propose, therefore, that a *Ki-ras* IVb coding

Calu-1 K-ras ATG ACT GAA TAT AAA CTT GTG GTA GTT GGA GCT TGT GGC GTA GGC AAG AGC GCC TTG ACG
 viral K-ras ATG ACT GAA TAT AAA CTT GTG GTA GTT GGA GCT TGT GGC GTA GGC AAG AGC GCC TTG ACG
 viral K-ras met thr glu tyr lys leu val val val gly ala asp gly val gly lys ser ala leu thr
 Calu-1 K-ras met thr glu tyr lys leu val val val gly ala oys gly val gly lys ser ala leu thr

Calu-1 K-ras ATA CAG CTA ATT CAO AAT CAT TTT GTC GAC GGA TAT GAT CCA AGA ATA GAG GAT TCC TAC
 viral K-ras ATA CAG CTA ATT CAA AAT CAT TTT GTC GAC GGA TAT GAT CCA AGA ATA CAG GAT TCC TAC
 viral K-ras ile gln leu ile gln asn his phe val asp gly tyr asp pro thr ile gln asp ser tyr
 Calu-1 K-ras ile gln leu ile gln asn his phe val asp gly tyr asp pro thr ile gln asp ser tyr

Calu-1 K-ras AGG AAG CAA GTA GTA ATT GAT GGA GAA ACC TOT CTC TTT GAT ATT CTC GAC ACA GCA GGT
 viral K-ras AGG AAG CAA GTA GTA ATT GAT GGA GAA ACC TOT CTC TTT GAT ATT CTC GAC ACA GCA GGT
 viral K-ras arg lys gln val val file asp gly glu thr oys leu leu asp ile leu asp thr ala gly
 Calu-1 K-ras arg lys gln val val file asp gly glu thr oys leu leu asp ile leu asp thr ala gly

Calu-1 K-ras CAA GAG GAG TAC AGT GCA ATG AGG GAC CAG TAC ATG AGC ACT GGG GAG GGC TTT CTT TGT
 viral K-ras CAA GAG GAG TAC AGT GCA ATG AGG GAC CAG TAC ATG AGC ACT GGG GAG GGC TTT CTT TGT
 viral K-ras val phe ala ile val file asp gly glu thr oys leu leu asp ile leu asp thr ala gly
 Calu-1 K-ras gln glu glu tyr ser ala met arg asp gln tyr met arg thr gly glu gly phe leu oys

Calu-1 K-ras GTA TTT GCC ATA AAT AAT ACT AAA TCA TTT GAA GAT ATT CAC CAT TAT AGA GAA CAA ATT
 viral K-ras GTA TTT GCC ATA AAT AAT ACT AAA TCA TTT GAA GAT ATT CAC CAT TAT AGA GAA CAA ATT
 viral K-ras val phe ala ile val file asp gly glu thr oys leu leu asp ile leu asp thr ala gly
 Calu-1 K-ras val phe ala ile val file asp gly glu thr oys leu leu asp ile leu asp thr ala gly

Calu-1 K-ras AAA AGA GAT AAG GAC TCT CAA GAT CCA CCT ATG GTC CTA GTA GGA AAT AAG TGT GAT TTT
 viral K-ras AAA AGA GAT AAG GAC TCT CAA GAT CCA CCT ATG GTC CTA GTA GGA AAT AAG TGT GAT TTT
 viral K-ras lys arg val lys asp ser glu asp val pro met val leu val gly asn lys oys asp leu
 Calu-1 K-ras lys arg val lys asp ser glu asp val pro met val leu val gly asn lys oys asp leu

Calu-1 K-ras CCT TCT ACA AGA CTA GAC AGA AAA CAG GCT CAG GAC TTA GCA AGA AGT TAT GGA ATT CCG
 viral K-ras CCT TCT ACA AGA CTA GAC AGA AAA CAG GCT CAG GAC TTA GCA AGA AGT TAT GGA ATT CCG
 viral K-ras pro ser arg thr val asp thr lys gln ala gln glu leu ala arg ser tyr gly ile pro
 Calu-1 K-ras pro ser arg thr val asp thr lys gln ala gln glu leu ala arg ser tyr gly ile pro

Calu-1 K-ras TTT ATG GAT AAG TCA GCA AAG ACA AGA CAG AGA GTC GAG GAT GCT TTT TAT ACA TTG GTC
 viral K-ras TTT ATG GAT AAG TCA GCA AAG ACA AGA CAG AGA GTC GAG GAT GCT TTT TAT ACA TTG GTC
 viral K-ras phe ile glu thr ser ala lys thr arg gln arg val glu asp ala phe tyr thr leu val
 Calu-1 K-ras phe ile glu thr ser ala lys thr arg gln arg val glu asp ala phe tyr thr leu val

Calu-1 K-ras AGA GAG ATC CGA CAA TAC AGA TTT AAA AAA ATC AGC AAA GAA GAA AAG ACT CCT GGC TGT
 viral K-ras AGA GAG ATC CGA CAA TAC AGA TTT AAA AAA ATC AGC AAA GAA GAA AAG ACT CCT GGC TGT
 viral K-ras arg glu ile arg gln tyr arg leu lys lys ile ser lys glu glu lys thr pro gly oys
 Calu-1 K-ras arg glu ile arg gln tyr arg leu lys lys ile ser lys glu glu lys thr pro gly oys

Calu-1 K-ras GGT AAA ATT AAA AAA TCG TTT ATA ATG TAA
 viral K-ras GGT AAA ATT AAA AAA TCG TTT ATA ATG TAA
 viral K-ras val lys ile lys lys oys val ile met TER
 Calu-1 K-ras val lys ile lys lys oys val ile met TER

Fig. 2 The complete nucleotide sequence and predicted amino acid sequence of the human Ki-ras gene (exons I, II, III and IVa) are shown together with the corresponding region of the viral Ki-ras gene determined by others¹⁰. Sequence differences are boxed. Arrowheads mark the splice sites of the human Ki-ras gene.

exon exists in rat with an amino acid sequence close to that shown in Fig. 3. We further propose that this sequence was transduced during the biogenesis of Ki-MuSV and, at some time during viral propagation or molecular cloning, an error in replication occurred, perhaps by polymerase slippage, in the dA-rich region which encodes a poly-lysine stretch. Since there has probably been no functional selection for the viral Ki-ras IVb sequence, it may also have diverged from the normal rat cellular sequence in other respects.

Transforming potential of Calu-1 Ki-ras gene

The Ki-ras gene of Calu-1, which transforms NIH 3T3 cells, differs from the Ki-ras gene of normal human DNA, which does not⁹. To determine the basis for this difference, we cloned a normal Ki-ras gene by screening a Charon 4A library constructed from human placental DNA. We obtained all but the 5'-most 3.7-kbp EcoRI fragment. The EcoRI restriction endonuclease map for the remaining 41.4 kbp was identical to that of the Calu-1 Ki-ras gene. It is possible that the difference between the normal and transforming Ki-ras genes results from small nucleotide changes. Indeed, the transforming potential of the normal Ha-ras gene can be activated by any mutation substituting valine⁶⁻⁸, arginine, aspartic acid or serine for glycine at amino acid position 12 (O.F., work in progress). Sequence comparisons of the first coding exons from the Calu-1 and normal Ki-ras gene showed two nucleotide differences. Significantly, Calu-1 Ki-ras encodes cysteine (TGT) at position 12 while the normal Ki-ras encodes glycine (GGT). Also, Calu-1 Ki-ras encodes glycine (GGA) at position 31 while the normal Ki-ras encodes glutamic acid (GAA) at that position. To test whether these sequence differences could account for the transforming capacity of the Calu-1 Ki-ras gene, the first coding exons of the Calu-1 and the normal Ki-ras genes were each joined by DNA ligation to the last three coding exons of a normal human Ha-ras gene. This construction used an XbaI site in the intron that separates the first from the second coding exons in both Ha-ras and Ki-ras genes (see Fig. 1). Only the chimaeric gene containing the first coding exon from Calu-1 Ki-ras was able to induce morphologically transformed foci of NIH 3T3 cells.

A
 a GGT GGT GAT GAT GAT GCC TTC TAT ACA TTA GAT CGA GAA ATT CGA AAA CAT AAA GAA AAG ATG AGC AAA GAT GGT AAA AAC AAG AAG AAG AAG TCA AAG ACA AAG TGT GTA AAT ATG TAA
 b GGT GGT GAT GAT GAT GCC TTC TAT ACA TTA GAT CGA GAA ATT CGA AAA CAT AAA GAA AAG ATG AAC AAA GAT GAG AAA AAA AAG AAG AAG AAG TCA AAG ACA AAG TGT GTA AAT ATG TAA
 AAAAA

B
 a gly val asp asp ala phe tyr thr leu val arg lys ile arg lys his lys glu lys met asn lys asp glu lys lys lys lys lys lys ser arg thr arg cys ile val met TER
 b gly val asp asp ala phe tyr thr leu val arg glu ile arg lys his lys glu lys met ser lys asp gly lys lys lys lys lys lys ser arg thr lys cys val file met TER
 c arg val glu asp ala phe tyr thr leu val arg glu ile arg gln tyr arg leu lys lys ile ser lys glu glu lys thr pro gly cys val lys ile lys lys cys val file met TER
 d arg val glu asp ala phe tyr thr leu val arg glu ile arg gln tyr arg leu lys lys ile ser lys glu glu lys thr pro gly cys val lys ile lys lys cys file ile met TER
 e gly val glu asp ala phe tyr thr leu val arg glu ile arg gln his lys leu arg lys leu asn pro pro asp glu ser gly pro gly cys met ser cys lys cys val leu ser TER
 f gly val glu asp ala phe tyr thr leu val arg glu ile arg gln his lys leu arg lys leu asn pro pro asp glu ser gly pro gly cys met ser cys lys cys val leu ser TER
 g gly val glu asp ala phe tyr thr leu val arg glu ile arg gln tyr arg met lys lys leu asn ser ser asp asp gly thr gln gly cys met gly leu pro cys val val met TER

Fig. 3 A Shows the complete nucleotide sequence of the human Ki-ras IVb exon (a) and its viral equivalent (b). Boxes indicate sequence differences. B, The predicted amino acid sequences of: a, rat Ki-ras exon IVb inferred from the Kirsten viral sequence¹⁰; b, human Ki-ras exon IVb; c, rat Ki-ras exon IVa inferred from the Kirsten viral sequence¹⁰; d, human Ki-ras exon IVa; e, rat Ha-ras exon IV inferred from the Harvey viral sequence²⁴; f, human Ha-ras exon IV^{11,12}; g, human N-ras exon IV (E.T., unpublished results). Differences between the human and rat homologues are boxed. Many of these differences are conservative (Arg for Lys or Ile for Val or Leu). Positions where the different members of the ras family contain the same or closely related amino acids are indicated by connecting lines. The sequence NBNNB, described in the text—a possible proteolytic cleavage sequence—is underlined.

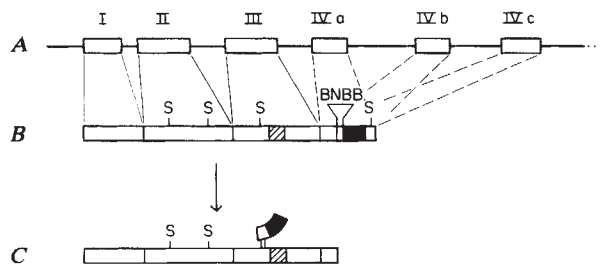


Fig. 4 Hypothetical structures of the *ras* genes and their products. **A**, The exonic arrangement of the *ras* genes with three coding exons (I, II and III) and alternate fourth coding exons (IVa, IVb, IVc and so on). The gene encodes a product, p21, 188–189 amino acids long (**B**) with domains encoded by the exons as indicated. The primary gene products have one highly variable region, indicated in black, and one less variable region, indicated by cross-hatching, which distinguish the individual members of the *ras* gene family but are conserved between species (rat and man). Preceding the highly variable region is the sequence BNBB (basic amino acid, any amino acid, basic, basic amino acid) which we postulate to be a proteolytic cleavage site. We suggest that in its mature form the terminal peptide containing the variable region is connected to the main body of the p21 through a disulphide linkage (||) to one of the cysteine residues indicated (S). Disulphide linkage of the type shown in **C** would bring the two variable regions into apposition, creating a physiologically important combining site unique for each member of the *ras* family.

Structure of *ras* genes and their products

A clearer picture is emerging of the relatedness of the *ras* gene products. First, there is strong conservation of the entire amino acid sequence between homologous *ras* genes in distant species. Thus, the human Ha-*ras* gene differs from the viral Ha-*ras* gene, derived from rat, at only 3 out of 189 amino acid positions, despite a high frequency of third-base neutral changes^{11,12}. Similarly, the human Ki-*ras* gene (comprised of exons I, II, III and IVa) differs from the viral Ki-*ras* gene, also derived from rat, at only 7 of 189 positions. We believe this conservation reflects evolutionary pressure on the maintenance of the physiological role of the *ras* gene products.

Second, the individual members of the *ras* gene family show a high degree of relatedness at the amino acid sequence level for amino acids 1–120, 133–163 and 185–189. In this 'constant' region the human Ki-*ras* and Ha-*ras* differ at only seven positions. The human N-*ras* gene shows a similar conservation in this region (E.T., unpublished results). However, the predicted amino acid sequence markedly diverges at amino acids 164–184, and less so at positions 121–132. We call these 'variable' regions (see Fig. 4). The sequence in the variable regions is, however, conserved between species. Thus there are no amino acid changes between the human and viral Ha-*ras* genes in this region and few significant differences between the human and viral Ki-*ras* genes. This region may, therefore, determine the physiological role of the individual *ras* gene products.

We wish to speculate further about the common structure of the *ras* gene products. In each case, the region of divergent amino acid sequence (171–184) begins with the amino acid sequence BNBB where B is either lysine or arginine and N is any amino acid. The dibasic amino acid sequence is a common recognition sequence for proteolytic cleavage of many membrane-associated and secreted proteins. The particular sequence BNBB is also commonly found at proteolytic cleavage sites, such as in preproinsulin^{16,17} and viral glycoprotein precursors¹⁸. Moreover, it is known from the work of others that the *ras* p21s undergo a carboxy-end proteolytic cleavage during maturation¹⁹. By comparing the electrophoretic mobility of synthesized, mature Ha-*ras* p21 with *in vitro*-translated Ha-*ras* p21 on reducing SDS-polyacrylamide gels, we estimate that the cleavage site is ~20 amino acids from the carboxy end. This would place cleavage quite near the BNBB sequence, separating the constant region from the variable region encoded by exon IV. What then is the fate of this variable region? One possibility is that after cleavage the terminal peptide serves as a hormone, growth factor or second messenger. Another possibility, the one we favour, is that this terminal peptide is retained by the *ras* p21 complex. Indeed, the last five amino acids of *ras* p21s are conserved and contain a cysteine residue, raising the possibility of a disulphide linkage of the terminal variable region with the main body of *ras* p21s. There are three conserved cysteine residues in the main body of the *ras* p21. One of these, encoded in the third exon, is immediately adjacent to another, less extensive, variable region. If disulphide linkage occurred here, a single, joint variable region would result (see Fig. 4). We propose that this region might serve as a combining site, allowing each different *ras* gene product to interact with a different cellular or extracellular constituent.

Finally, we wish to speculate about the structure of the *ras* genes themselves. The exonic structures of the human Ha-*ras* and Ki-*ras* genes are quite similar, indicating that both have evolved from a similarly spliced ancestral gene. However, an alternative fourth coding exon was found in the Ki-*ras* gene by virtue of the presence of homologous sequences in the Ki-MuSV genome. Further alternate fourth coding exons may be present, not only in the Ki-*ras* gene but in the other *ras* genes also. These might not be readily detected by hybridization analysis if they are as diverged as Ki-*ras* IVa is from Ki-*ras* IVb. Thus, not only are there at least five members of the *ras* gene family^{3,20}, but each member may be able to express more than one physiologically distinct protein. The control of the expression of these genes and their gene products is probably prominent in growth control, development and associated pathology.

This work was supported by grants from the NCI (CA 29569), the Robertson Research Fund, and the American Business for Cancer Research Foundation. Y.S. is the recipient of a postdoctoral fellowship from the Schweizerische Stiftung für Medizinisch-Biologische Stipendien. D.B. is supported by Institut National de la Santé et de la Recherche Médicale and Philippe Foundation Fellowships. E.T. is supported by a NIH Postdoctoral Fellowship.

Received 16 May; accepted 27 May 1983.

1. Ellis, R. *et al.* *Nature* **292**, 506–511 (1981).
2. Der, C., Krontiris, T. & Cooper, G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637–3640 (1982).
3. Parada, L., Tabin, C., Shih, C. & Weinberg, R. *Nature* **297**, 474–478 (1982).
4. Santos, E., Tronick, S., Aaronson, S., Pulciani, S. & Barbacid, M. *Nature* **298**, 343–347 (1982).
5. Shimizu, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 2112–2116 (1983).
6. Taparowsky, E. *et al.* *Nature* **300**, 762–765 (1982).
7. Reddy, E., Reynolds, R., Santos, E. & Barbacid, M. *Nature* **300**, 149–152 (1982).
8. Tabin, C. *et al.* *Nature* **300**, 143–148 (1982).
9. Perucho, M. *et al.* *Cell* **27**, 467–476 (1981).
10. Tsuchida, N., Ryder, T. & Ohtsubo, E. *Science* **217**, 937–939 (1982).
11. Fasano, O., Taparowsky, E., Fiddes, J., Wigler, M. & Goldfarb, M. *J. molec. appl. Genet.* (in the press).

12. Capon, D., Chen, E., Levinson, A., Seeburg, P. & Goeddel, D. *Nature* **302**, 33–37 (1983).
13. Ait, F. *et al.* *Cell* **20**, 293–301 (1980).
14. Early, P. *et al.* *Cell* **20**, 313–319 (1980).
15. Amara, S., Jonas, V., Rosenfeld, M., Ong, E. & Evans, R. *Nature* **298**, 240–244 (1982).
16. Ullrich, A. *et al.* *Science* **196**, 1313–1319 (1977).
17. Perler, F. *et al.* *Cell* **20**, 555–566 (1980).
18. Rice, C. & Strauss, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2062–2066 (1981).
19. Shih, T. *et al.* *J. Virol.* **42**, 253–261 (1982).
20. Chang, E., Gonda, M., Ellis, R., Scolnick, E. & Lowy, D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4848–4852 (1982).
21. Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
22. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–580 (1980).
23. Drouin, J. *J. molec. Biol.* **140**, 15–34 (1980).
24. Dahr, R. *et al.* *Science* **217**, 934–937 (1982).