

Sequence of cDNA coding for human keratin 19

Pauline C.Stasiak and E.Birgitte Lane

Cell Structure Laboratory, Imperial Cancer Research Fund Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, UK

Submitted October 19, 1987

Accession no.Y00503

We present the cDNA sequence of human keratin 19 (40kD;(1)) determined from clones isolated from a lambda gt11 expression library of human placenta by using specific monoclonal antibodies. The sequence has a full-length open reading frame (initiating ATG and stop codon are underlined). Although keratin 19 is considered characteristic of simple epithelia, the sequence is clearly closer to human keratinocyte keratins (2,3) than to the simple epithelial keratin 18 (4). It is very similar to that of the homologous bovine keratin (5): 73% bases are identical in the 3' untranslated region, and in the coding region there is 86% identity over the head domain and 89.5% over the helical rod domain. In the short tail domain (pos'ns. 1168-1207) this falls to only 64%, even though other keratins maintain over 80% identity with homologous proteins across species boundaries in all three domains. Translation of this sequence shows the tail homology to be even lower at the amino acid level (38%: only 5 out of the 13 amino acids shared). This high divergence between human and bovine sequences over the tail domain suggests that little if any selection pressure has been acting on this region. This indicates that the abnormally short tail of keratin 19 is, like our own, vestigial.

```

CCGCTCGCC ATG ACT TCC TAC AGC TAT CGC CAG TCG TCG GCC ACG TCG TCC TTC GGA GGC      61
CTG GGC GGC GGC TCC GTG CGT TTT GGG CCG GGG GTC GCT TTT CGC GCG CCC AGC ATT CAC
GGG GGC TCC GGC CGC GGC GGC GTA TCC GTG TCC TCC GCC CGC TTT GTG TCC TCG TCC TCC      181
TCG GGG GGC TAC GGC GGC GGC TAC GGC GGC GTC CTG ACC GCG TCC GAC GGG CTG CAC GGC
AAC GAG AAG CTA ACC ATG CAG AAC CTC AAC GAC CGC CTG GCC TCC TAC CTG GAC AAG GTG      301
CGC GGC CTG GAG GCG GCC AAC GGC GAG CTA GAG GTG AAG ATC CGC GAC TGG TAC CAG AAG
CAG GGG CCT GGG CCC TCC CGC GAC TAC AGC CAC TAC TAC ACG ACC ATC CAG GAC CTG CCG      421
GAC AAG ATT CTT GGT GCC ACC ATT GAG AAC TCC AGG ATT GTC CTG CAG ATC GAC AAC GCC
CGT CTG GCT GCA GAT GAC TTC CGA ACC AAG TTT GAG ACG GAA GAG GCT CTG CGC ATG AGC      541
GTG GAG GCC GAC ATC AAC GGC CTG CGC AGG GTG CTG GAT GAG CTG ACC CTG GCC AGG ACC
GAC CTG GAG ATG CAG ATC GAA GGC CTG AAG GAA GAG CTG GCC TAC CTG AAG AAG AAC CAT      661
GAG GAG GAA ATC AGT ACG CTG AGG GGC CAA GTG GGA GGC CAG GTC AGT GTG GAG CTG GAT
TCC GCT CCG GGC ACC GAT CTC GCC AAG ATC CTG AGT GAC ATG CGA AGC CAA TAT GAG GTC      781
ATG GCC GAG CAG AAC CCG AAG GAT GCT GAA GCC TGG TTC ACC AGC CGG ACT GAA GAA TTG
AAC CCG GAG GTC GCT GGC CAC ACG GAG CAG CTC CAG ATG AGC AGG TCC GAG GTT ACT GAC      901
CTG CCG CGC ACC CTT CAG GGT CTT GAG ATT GAG CTG CAG TCA CAG CTG AGC ATG AAA GCT
GCC TTG GAA GAC ACA CTG GCA GAA ACG GAG CCG CGC TTT GGA GCC CAG CTG GCG CAT ATC      1021
CAG CCG CTG TAC AGC GGT ATT GAA GCC CAG CTG GCG GAT GTG CGA GCT GAT AGT GAG CCG
CAG AAT CAG GAG TAC CAG CCG CTC ATG GAC ATC AAG TCG CCG CTG GAG CAG GAG ATT GCG      1141
ACC TAC CGC AGC CTG CTC GAG GGA GAG GAA GAT CAC TAC AAC AAT TTG TCT GCC TCC AAG
GTC CTC TGA GGCAGCGGCTCTGGGGCTTCCTGCTGCTCTTTGGAGGGTGTCTTCTGGGTAGAGGGATGGGAAGGAA      1277
GGGACCCCTTACCCCGGGCTCTTCTCTGACCTGCCAATAAAAAATTTATGGTCCAAGGGAAAAAATAAAAAAAAAA
AAAAAAAAAAAAA
                                                                                               1369
    
```

REFERENCES 1.Wu J-Y and Rheinwald J.G. (1981) Cell 25 627-635. 2.Hanukoglu I. and Fuchs E. (1982) Cell 31 243-252. 3. RayChaudury A., Marchuk D., Lindhurst M., Fuchs E. (1986) Mol. and Cell Biol. 6 539-548. 4. Oshima R.G., Millan J.L., Cecena G. (1986) Differentiation 33 61-68. 5. Bader B.L., Magin T.M., Hatzfeld M., Franke W.W. (1986) EMBO J. 5 1865-1875.