B 3139

A kapcsoló fehérje molekuláris evolúciójának és a gén 5' végének vizsgálata

Doktori dolgozat

BARTA ENDRE

MTA Szegedi Biológiai Központ
Biokémiai Intézete
SZEGED

1993

KÖSZÖNETNYILVÁNÍTÁS

Köszönettel tartozom Dr. Duda Ernőnek, aki pályámon elindított, és megismertetett ennek a munkának a szépségeivel.

Nagyon sok hálával tartozok témavezetőmnek, Dr. Deák Ferencnek, akitől legtöbbet tanulhattam eddigi munkám során. Mindig számíthattam a segítségére, bármilyen jellegű problémáról volt is szó.

Nagyon sok köszönettel tartozom csoportvezetőnknek, Dr Kiss Ibolyának, aki végig figyelemmel kísérte a munkámat és észrevételeivel, kritikai megjegyzéseivel segítette azt. Köszönöm neki és Dr. Deák Ferencnek, hogy mindvégig bátorítottak ennek a dolgozatnak a megírására és nagyon sokat segítettek még akkor is amikor már nem dolgoztam a csoportjukban.

Köszönöm munkatársaimnak, hogy ötleteikkel, észrevételeikkel segítették munkámat, vagy egyes kísérleteimben tevékenyen is részt vállaltak. Nevük álljon itt időrendi sorrendben: Silvija Mestric, Markus Biesold, és Bachrati Csanád.

Köszönettel tartozom Dr. Fekete Endréné asszisztensnek és Takács Gábornak aki a fotókat készítette.

Hálás köszönettel tartozom a Gödöllői Mezőgazdasági Biotechnológiai Központból Dr. Balázs Ervinnek és Dr. Bánfalvi Zsófiának, akik új munkahelyemen is lehetőséget biztosítottak kísérleteim befejezésére és dolgozatom megírására.

Dolgozatom megírását segítő kritikai észrevételeit köszönöm Dr. Szabó Gábornak és Dr. Szekeres Miklósnak.

Végül, de nem utolsósorban tartozom köszönettel feleségemnek Fazakas Andreának, akinek a segítségével kerültem először kapcsolatba a Kötőszövet Laboratóriummal, és aki később is amikor már nem dolgozott ott, figyelemmel kísérte munkámat és türelmével, megértésével rendkívül sokat tett azért, hogy ez a dolgozat elkészülhessen.

TARTALOMJEGYZÉK

		oldal
RÖVIDÍTÉSEK		1.
BEVEZETÉS		2.
IRODALMI ÁTTEKINTÉS		3.
I.	A porcszövet	3.
II.	A kapcsoló fehérje	7.
	1. A kapcsoló fehérje felfedezése	7.
	2. A kapcsoló fehérje gén	8.
III.	A molekuláris evolúció kérdései	10.
	1. Az intronok evolúciós szerepe	10.
	2. Mozaik fehérjék, exon shuffling	11.
	3. A homológ modulok jelentősége és vizsgálata	12.
IV.	A kapcsoló fehérjével homológ modulokat tartalmazó fehérjék	13.
	1. Az első hurokkal rokon modulok	13.
	2. A második és harmadik hurokkal rokon modulok	14.
CÉLKITŰZÉS		17.
ANYAGOK ÉS MÓDSZEREK		
I.	Kísérleti anyagok, baktérium törzsek, plazmidok, fágok	18.
	1. Vegyszerek	18.
	2. Törzsek, vektorok	19.
	3. Táptalajok	19.
	4. Oldatok	19.
II.	Módszerek	20.
	1. Plazmid DNS izolálás	20.
	2. M13 fág DNS izolálás	20.

3. Lambda fág DNS izolálás 21.
4. Restrikciós emésztés 22.
5. DNS fragment izolálás 22.
6. Rádioaktív próbák jelölése 22.
7. DNS elektroforézis 22.
8. RNS izolálás 22.
9. Northern hibridizáció 23.
10. Primer extenzió 23.
11. cDNS klónozás 23.
12. Telephibridizáció 23.
13. Southern transzfer és hibridizáció 24.
14. Szekvenálás 25.
15. Polimeráz láncreakció (PCR) 25.
16. T4 DNS polimeráz térképezés 26.
17. Komputeres analízis 26.

EREDMÉNYEK 27.

I. cDNS klónozás 27.
1. A mRNS tisztítása 27.
2. A primer extenzió 27.
3. A klónozás menete 28.

II. cDNS klónok kiválasztása és jellemzése 30.
1. A kapcsoló fehérje mRNS-sel komplementer klónok kiválasztása 30.
2. A cDNS klónok restrikciós térképezése 32.
3. Az izolált cDNS klónok szekvenciájának meghatározása 34.

III. Alternatív exon keresés a második és a harmadik exonok között 34.
1. A pozitív cDNS klónok közötti keresés 34.
2. Polimeráz láncreakció az alternatív exon kimutatására 34.

IV. Az első exont tartalmazógenomiális klón 35.
1. A λgLP100 klón térképezése 35.
2. A pLPD1 klón 38.

V. Alternatív splicing a kapcsoló fehérje mRNS 5' végi nem transzlálódó részén 40.
1. Konszenzus kivágódási helyek megléte 42.
2. Alternatív splicing bizonyítása polimeráz láncreakcióval 42.

VI. A transzkripció kezdőpontjának meghatározása 44.
1. Primer extenzió a PEr és a PFr primerekkel 44.

2. Transzkripciós startpont meghatározása T4 polimerázzal 46.

VII. A genomiális DNS szekvencia számítógépes elemzése 47.

VIII. A kapcsoló fehérje gén molekuláris evolúciója 48.
1. Az első hurok 48.
2. A második hurok vizsgálata 51.
3. A kapcsoló fehérje molekuláris evolúciója 53.

EREDMÉNYEK MEGVITATÁSA 56.

I. A kapcsoló fehérje gén 5' vége 56.

II. A kapcsoló fehérje evolúciója 63.

ÖSSZEFOGLALÁS 66.

A DOLGOZAT ANYAGÁBÓL KÉSZÜLT PUBLIKÁCIÓK 68.

IRODALOM 69.

RÖVIDÍTÉSEK

bp	:bázispár
CMP	:porc mátrix fehérje
CS	:kondroitin-szulfát
DTT	:ditiotreitol
EDTA	:etilén-diamin-tetraecetsav
Ig	:Immunglobulin
IPTG	:izopropil-tio-β-D-galaktozid
kb	:kilobázispár
kDa	:kilodalton
KS	:keratán-szulfát
MES	:2-[N-morfolino]-etán-szulfonsav
MOPS	:3-[N-morfolino]-propán-szulfonsav
PCR	:polimeráz láncreakció
PTR	:"proteoglycan tandem repeat", a kapcsoló fehérje ismétlődő egységével rokon modulok
rpm	:fordulat per perc
SDS	:nátrium-dodecil-szulfát
TNF	:tumor nekrózis faktor
Tris	:trisz-/hidroxi-metil/-amino-metán
X-gal	:5-bromo-4-kloro-3-indolil-β-D-galaktozid

GHAP

SSC

dNTP

TSG-6

BEVEZETÉS

Az élővilág egyre mélyebb megismerésében fontos terület az eukarióta génreguláció tanulmányozása. A különböző gének eltérő időben történő be- és kikapcsolása, különböző szinten történő expressziója, és a hírvívők átírása fehérjévé határozza meg az egyedfejlődés menetét. Ennek a sokrétű, bonyolult folyamatnak a megismerését segíti a különböző modellrendszerek tanulmányozása. A gerinces állatok esetében a génreguláció vizsgálatához nagyon jó modell a porcszövet, mivel rendkívül specializált, viszonylag kevés gén fejeződik ki benne az egyedfejlődés egy meghatározott rövid szakaszában, és csak a szervezet bizonyos helyein.

1987 és 1991 között alkalmam volt bekapcsolódni a Szegedi Biológiai Központ Biokémiai Intézetében a kötőszövet molekuláris biológiáját kutató csoport munkájába. A dolgozatban leírt eredmények az itt végzett kollektív munkán alapulnak. Amennyiben valamelyik eredmény nem a saját munkámon alapszik azt külön jeleztem. A csoport két fehérje gén vizsgálatával foglalkozik, amelyek a csirke porc mátrix fehérjét (CMP) és a csirke kapcsoló fehérjét kódolják. Mindkét génnek izolálták a cDNS és genomiális klónjait, valamint a CMP gén promoter régióját. Én a kapcsoló fehérje mRNS 5' végének és a gén promoter régiójának az izolálásában vettem részt, ezt a munkát tartalmazza a dolgozat első fele.

A kapcsoló fehérje aminosavsorrendjének és a gén exon-intron szerkezetének a megismerése felfedte, hogy a fehérje, sok más extracelluláris fehérjével együtt, moduláris szerkezetű, ami azt jelenti, hogy a különböző domének az evolúció során külön keletkeztek. A dolgozat második felében a kapcsoló fehérje molekuláris evolúcióját vizsgálom a különböző modulok számítógépes elemzésével.

IRODALMI ÁTTEKINTÉS

I. A porcszövet

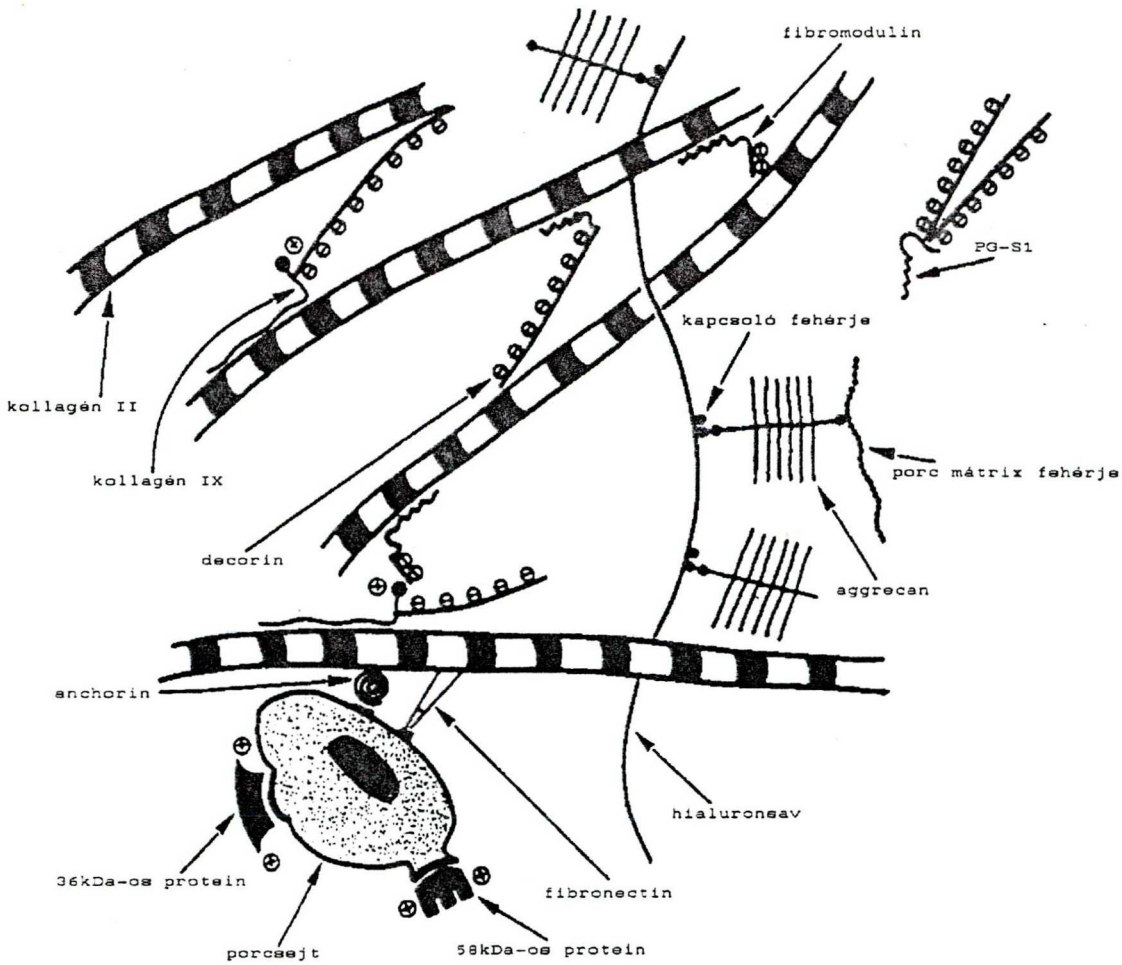
A porcszövet a gerincesekre jellemző támasztószövet, amely elsősorban sejtközötti állománya miatt nagyon rugalmas, ugyanakkor igen nagy összenyomási szilárdságú. A porc alkotja a porcos halak, valamint a magasabb rendű gerincesek embrióinak vázát. Az egyedfejlődés során azonban a magasabbrendű gerincesekben helyét a sokkal szilárdabb csont veszi át, a fej csontjai kivételével az összes csontra a porcos csontosodás jellemző. A porcszövet a csontok hosszirányú megnyúlását biztosító növekedési korongokban, valamint a csontok végeinél található ízületekben marad meg, ahol elsősorban a csontok sima elmozdulását biztosítja. A porcszövet érdekessége, hogy a többi szövetféleséggel szemben benne sem idegek, idegvégződések, sem vérerek nem találhatók. A táplálékok, szabályozó anyagok és az oxigén a sejtközötti állományon keresztül diffúzióval jut el a porcsejtekhez (1-5).

A kifejlett madarakban, emlősökben a porcszövetet tulajdonságai és előfordulási helye szerint három csoportba lehet osztani, úgymint

- hyalin, vagy üvegporc, amely megtalálható a különféle nagy összenyomási terhelésnek kitett helyeken, például az ízületekben,
- elasztikus vagy rugalmas porc, amely olyan helyeken fordul elő, mint a fülkagyló vagy a gégefedő, rendkívül ellenálló hajlításoknak és alapállományának jelentős részét a rugalmas, elasztikus rostok teszik ki,
- kollagénrostos porc, amely az erősebb szakító-, nyíró-, és csavaró behatásomxak kitett helyeken található, mint a csigolya közti, vagy az ízületi közbeiktatott porckorongok.

Ezek közül a legjobban tanulmányozott porcfajta az üvegporc. Speciális, fizikai és kémiai behatásoknak renkívül ellenálló felépítettsége miatt az üvegporc beható vizsgálata akkor vált lehetővé, amikor először alkalmazták a guanidium-kloridot a proteoglükánok kivonására (6), valamint tovább finomították a különböző proteoglükánok elválasztását lehetővé tevő cézium-klorid egyensúlyi sűrűség grádiens centrifugálást (7,8). E módszerek alkalmazásával sikerült izolálni és karakterizálni a porc mátrix makromolekuláit.

Az üvegporcról alkotott nézeteink elsősorban a csirke embriók mellcsont kezdeményének, a csirke mellcsont porcos nyúlványának (xyphisternum), a szarvasmarha orrsövényének és légcsövének, az ember ízületi felszíneit (elsősorban a térdízületet) beborító porcnak, valamint a Swarm-féle patkány kondroszarkomának (9) a vizsgálatából származnak.



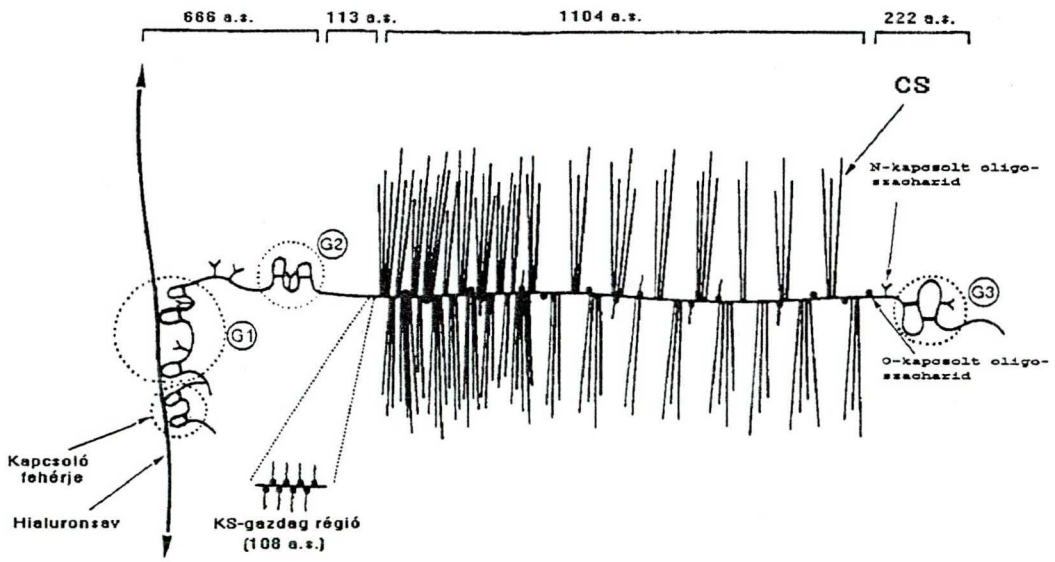
1. ábra: A porc mátrix jellegzetes makromolekulái és a különböző makromolekulák egymással, illetve a porcsejttel alkotott kölcsönhatásai.

A porcszövetet az egyesével vagy kis csoportokban elhelyezkedő porcsejtek, és a közöttük elhelyezkedő igen kiterjedt sejtközötti állomány alkotja (1. ábra). A sejtközötti állomány a szövetben szórványosan elhelyezkedő porcsejtek terméke. A sejtközötti állomány fehérjéit három csoportba lehet sorolni, úgymint kollagének, proteoglükánok, és a nem kollagénszerű fehérjék.

A kollagének általánosan előfordulnak a kötőszövetekben, hálózatot alkotnak, amely körülveszi a sejteket és amelyekbe beleágyazódnak a sejtközötti állomány többi molekulái (10). Általános jellemzőjük, hogy rostokat, hálózatokat alkotnak, melyek alapegységei a tripla hélixek (11). A porcban a kollagén állomány nagy részét a II. típusú kollagén alkotja (12). A II. típusú kollagén három azonos, telopeptidekkel határolt, 1014 aminosavat tartalmazó, balmenetes hélixet képező α láncból áll. A csontban előforduló kollagénekhez képest rostjai fi-

nomabbak, és hálózatokat alkotnak párhuzamos kötegek helyett. A feladatuk is eltérő. A csontban a kollagének felelősek a feszítéssel és hajlítással szembeni ellenállásért, míg a porcban a kollagénrostok elosztják a terhelést, és "lekötözik" a proteoglükán aggregátumokat. A porcszövet kollagénjeinek 5%-át teszi ki a IX. típusú kollagén (13, 14), amely kovalensen kötődik a II. típusú kollagénhez (15), de megtalálható a X. és a XI. típusú kollagén is. A X. típusú kollagén az elcsontosodó porc fő komponense.

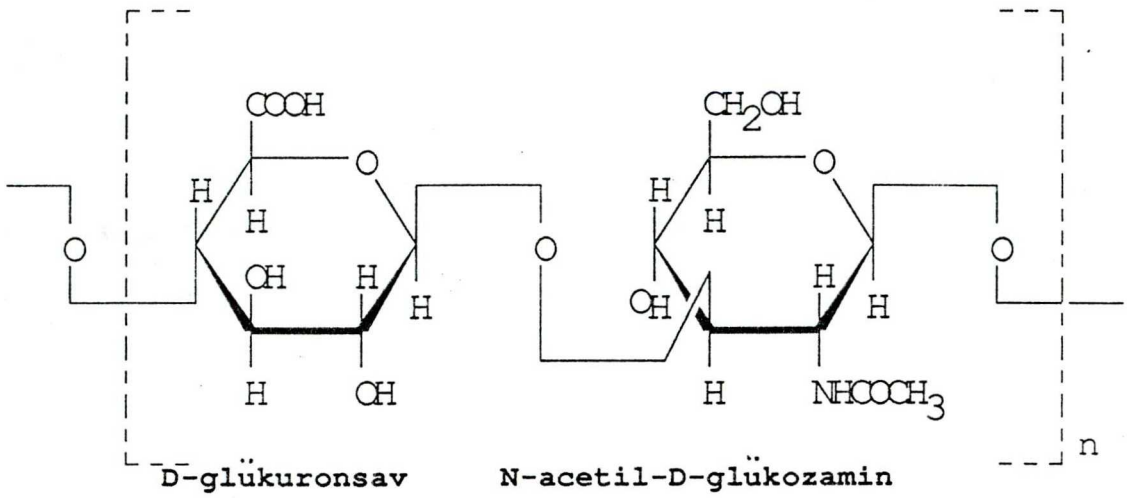
A proteoglükánokra az jellemző, hogy egy központi polipeptidlánchoz különböző glükózaminoglükán oldalláncok kapcsolódnak (16, 17). Az üvegporc nedves súlyának mintegy 5-10%-át teszi ki az ún. nagy aggregálódó, vagy másnéven kondroitin-szulfát oldal-



2. ábra: A patkány kondroszarkoma aggrecan szerkezete, a hialuronsavval és a kapcsoló fehérjével való kapcsolódása. a.s. :aminosav, KS: keratán-szulfát, CS: kondroitin-szulfát.

láncú proteoglükán, az aggrecan (2. ábra). A patkány kondroszarkoma aggrecan molekula 2124 aminosavból áll, ami 207.8 kDa molekulatömeget jelent (18). Az aggrecan molekula N-terminális felén két globuláris egység található (G1, G2), a C-terminális végén pedig egy harmadik globuláris egység (G3) helyezkedik el, amely a lektinekkel mutat hasonlóságot (19). A polipeptid gerincen 97 kondroitin-szulfát, 30 keratán-szulfát, és 42 heparán-szulfát oldallánc, valamint számos kovalensen kötött oligoszacharid lehetséges kapcsolódási helye található (19), jó egyezésben a korábbi becslésekkel (20). A fehérje molekula gerincén oldalirányban elhelyezkedő glükozaminoglükán oldalláncokon található erősen negatív töltésű szulfát és karboxil csoportok miatt az aggrecan híg oldatban rendkívül nagy térkitöltésű (21). Ugyanakkor a porcszövet extracelluláris mátrixában sokkal jobban össze van

nyomva, hiszen a térfogata az ötödére csökken többek közt a kollagén hálózat jelenléte miatt. Ebben az állapotban az oldalláncok nagyon közel kerülnek egymáshoz, így a köztük lévő elektrosztatikus taszítóerő is magas lesz. Ez hozzájárul a porcszövet fizikai tulajdonságainak, a nagy rugalmasságnak, a fizikai hatásokkal szembeni nagyfokú rezisztenciának a kialakulásához. Az aggrecan molekula jellegzetessége, hogy monomerjei hialuronsav jelenlétében aggregálódnak (22).



3. ábra: A hialuronsav makromolekula kémiai felépítése

A hialuronsav (3. ábra) ismétlődő D-glükuronsav és N-acetil-D-glükozamin diszacharid egységekből álló mukopoliszacharid. Megtalálható a sejteket borító rétegekben, a gerincesek kötőszövetében mint extracelluláris alapanyag, továbbá az ízületek porcaiban. Kémiai felépítése és hosszú láncú molekulája érthetővé teszi, hogy ez a nagy molekulatömegű poláros vegyület a vízzel rendkívül viszkózus anyagot ad, ami hozzájárul a porc rugalmasságának a kialakításához is.

Az aggrecanon kívül a porcszövetben előfordulnak még más kisebb molekulatömegű, vagy nem aggregálódó proteoglükánok is. A biglükán vagy PG-S1 fehérjelánc 38 000, míg a vele részben homológ decorin vagy PG-S2 36 500 dalton molekulatömegű. Mindkét molekula szöveti eredettől függően tartalmazhat kondroitin-szulfát, vagy dermatán-szulfát oldalláncot. A porcban azonos moláris mennyiségben fordulnak elő az aggrecannal, kis molekulatömegük miatt azonban a teljes tömeg mindössze 1-2 százalékát képviselik. Ezzel szemben a rostos kötőszövetben a proteoglükánok 95 százalékát a kis molekulatömegű proteoglükánok képviselik (16, 23, 24).

A nem kollagénszerű fehérjék közül legjobban ismert kapcsoló fehérjéről egy későbbi fejezetben lesz szó részletesen. A porc matrix fehérjéről annyit tudunk, hogy előfordulása a porcszövetekre korlátozódik, bár leírták a szem bizonyos rétegeiben is (25). Valószínűleg kapcsolódik mind az aggrecanhoz, mind a II. típusú kollagénhez, így a porc sejtközötti állományának a stabilizálódásában játszthat szerepet. A csirke porc mátrix fehérje prekurzora 493 aminósav hosszú, 54 kDa makromolekula (26). A kapcsoló fehérjéhez és az aggrecanhoz hasonlóan a porc mátrix fehérje is mozaik szerkezetű, két ismétlődő von Willebrand faktor A modult tartalmaz, melyek egy epidermális növekedési faktor modult fognak közre (27). Ismert a gén promotere, továbbá a gén expressziójának szabályozásában résztvevő pozitív és negatív régiók elhelyezkedése. A gén első intronjában egy szövetspecifikus enhanszer található (28).

Az egyéb, a porc extracelluláris mátrixából leírt fehérjék közül a fibromodulin, vagy 59-kDa protein 375 aminosav hosszú, erős homológiát mutat a decorinnal és a biglükánnal. Kötődik a kollagénhez, és in vitro befolyásolja a kollagénrostok kialakulását (16, 29). A porcsejtek felszínéhez tapad, vagy valamilyen módon kölcsönhatásba lép velük a 36-, az 58-kDa protein, az anchorin cII, valmint a fibronektin (16). Ennek funkciójáról azonban jelenleg még nagyon keveset tudunk.

II. A kapcsoló fehérje

1. A kapcsoló fehérje felfedezése

Hascall és Sajdera (7) mutatták ki, hogy a marha orrporcból izolált proteoglükán aggregálódni képes egy úgynevezett "protein-link" frakció jelenlétében. Később kiderült (22), hogy ebben a "protein-link" frakcióban az aggregálódásért inkább a hialuronsav felelős, de a hialuronsav mellett található még a frakcióban néhány fehérje komponens is, melyek szintén részt vehetnek az aggregálódásban (20). Ezeket a fehérjéket poliakrilamid gélen elválasztva két fő komponenst találtak, melyeket elneveztek kapcsoló fehérje 1 és 2-nek (30). Később még egy komponenst ki tudtak mutatni, a kapcsoló fehérje 3-mat. E három fehérjének a molekulatömege marha porcban 51, 46 és 41 KDa (31). A két forma proteolítikus hasítása azt mutatja, hogy apolipeptid gerincük megegyezik, de míg az 1-es forma 9.5%, a 2-es csak 3% szénhidrátot tartalmaz az aminosav oldalláncokhoz kötve. A különbségek az N terminális régióra lokalizálódnak (32, 33).

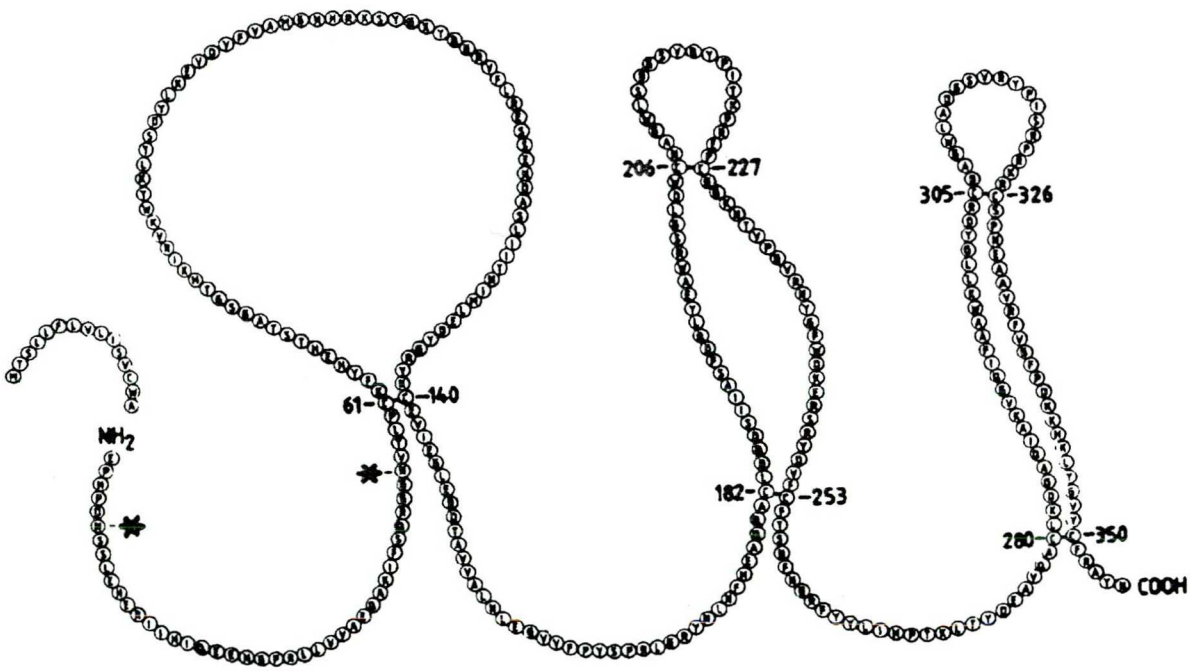
Mint már említettük, a kapcsoló fehérje részt vesz a proteoglükán-hialuronsav aggregátumok képzésében, azokat stabilizálja (34-36). Erre azáltal képes, hogy kötődik mind az aggrecanhoz (37, 38), mind a hialuronsavhoz (39, 40). A proteoglükán-hialuronsav

aggregátumok hialuronidázzal történő emésztéséből kiderült, hogy az aggregatum 1 molekula aggrecant, 1 molekula kapcsoló fehérjét, és átlagosan 41 monoszacharid egységből álló hialuronsav oligoszacharid egységet tartalmaz (41).

A kapcsoló fehérjét a porcon kívül, különböző immunohisztokémiai módszerekkel kimutatták még az aortában (42, 43), a szem inhártyájában (44) és a synoviális sejtekben (45).

2. A kapcsoló fehérje gén

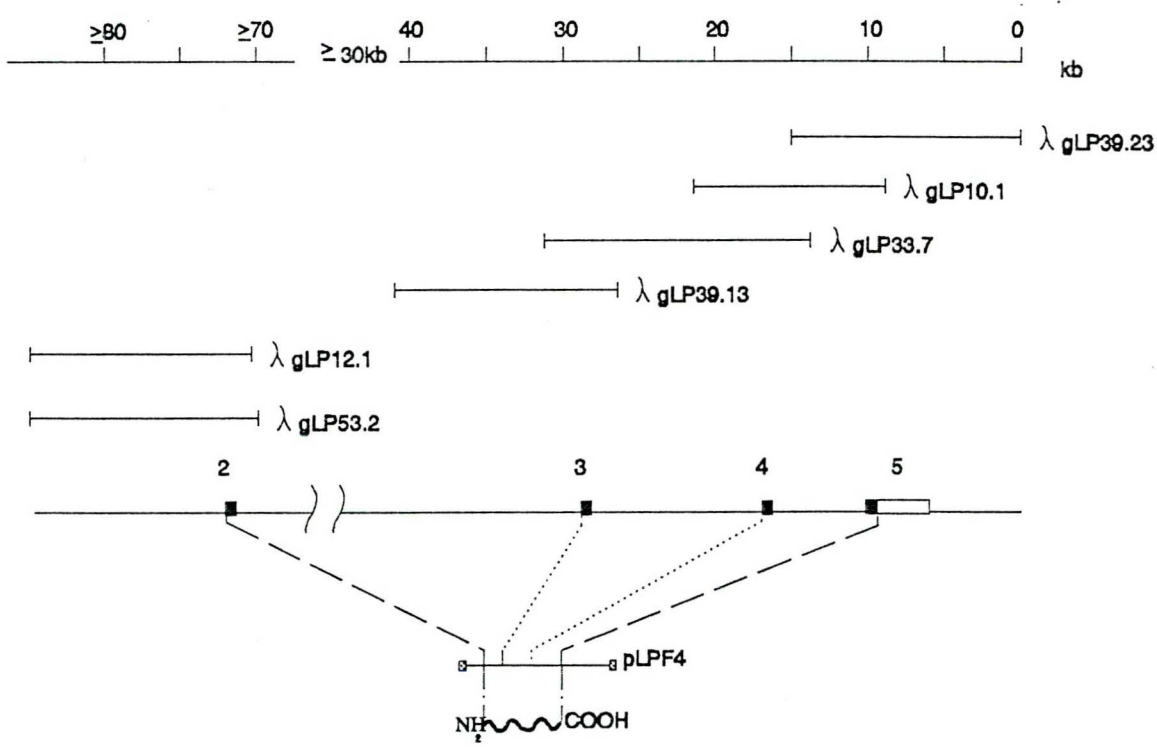
A kapcsoló fehérje teljes kódoló régióját hordozó cDNS-t először Deák F. és mti. izolálták csirkéből 1986-ban (46). Rövid peptiddarabok szekvenciája ismert volt marha porc-



4. ábra: A kapcsoló fehérje aminosavsorrendje a diszulfid hidak helyzetével és pozíciójával. A lehetséges N-glikozilációs helyeket csillagok jelzik.

ból (47) és patkány kondroszarkóma (48) sejtekből. Kevert oligonukleotid szintézissel ezeket kódoló oligonukleotidokat szintetizáltattak, és ezekkel 14 napos csirke embriók porcos mellcsont kezdeményéből készített cDNS könyvtárat vizsgáltak át. Találtak is egy klónt, amely specifikus volt a porcra, és ennek a klónnak a segítségével további klónokat sikerült izolálni a könyvtárból. Meghatározták az átfedő cDNS klónok restrikciós térképét, és a közös régió szekvenciáját. Egy nagy nyitott leolvasási keretet találtak, amely 1065 bp hosszú és 365 aminosavat határoz meg (4. ábra). Ennek a fehérjének a számolt molekulatömege 41.7 kDa, ami jól megegyezik az irodalomban a fehérje nem glükozilált formájára közölt 41 kDa-os

értékkel. Northern hibridizációs kísérletek szerint háromféle kapcsoló fehérje specifikus mRNS található 6.0, 5.8 és 3.0 kb hosszúságban. A különbségek a 3' végből adódnak. Neame és mti. időközben meghatározták a patkány kondroszarkóma kapcsoló fehérje különféle proteázokkal való emésztésével nyert fragmentumainak aminosavsorrendjét, és a diszulfid hidak helyzetét (50). Doege és munkatársai pedig egy cDNS klónt találtak, amely a patkány kapcsoló fehérje kétharmadát fedte le (51). Mára ugyancsak ismert a kapcsoló fehérje cDNS szekvenciája emberből, (52, 53) és sertésből is (52).



5. ábra: A kapcsoló fehérje gén szerkezete a pLPF4 cDNS klón és 6 különböző genomiális klón alapján (54)

A cDNS klónok segítségével sikerült meghatározni a csirke kapcsoló fehérje gén exon-intron szerkezetét is (54). A gén fehérjét kódoló része 4 exonra tagolódik (5. ábra), amelyek együttesen a genom 80 kb-t meghaladó hosszúságú részét foglalják el. A DNS-RNS hibridek elektromikroszkópos vizsgálata arra utal, hogy a gén 5' végén van még egy exon, de ezt az izolált genomiális klónok még nem tartalmazták. Ezt a megfigyelést megerősítette, hogy a klónok közül a leghosszabb, a pLPF4 109 bp-ral túlnyúlt 5' irányban a klónozott exonokon.

Ugyancsak ismert a patkány kapcsoló fehérje exon-intron struktúrája (55). Az intron kivágódási helyek hasonló pozíciókban helyezkednek el, mint a csirkében. A mRNS-ek 10%-ánál azt találták, hogy a 3. exon előtt egy új exon található. Ez az alternatív exon, amennyiben átíródik fehérjévé, alapjában megváltoztathatja a fehérje tulajdonságait, mivel erősen pozitívan töltött (53 aminosavban 5 lizin, és 5 arginin található), és három ciszteint is tartalmaz. Mindenesetre in vivo bizonyíték még nincs ennek a fehérjeterméknek a létezésére.

III. A molekuláris evolúció kérdései

1. Az intronok evolúciós szerepe.

A hetvenes évek végére világossá vált, hogy az eukarióta gének azon régiói, amelyek az érett mRNS-t alkotják, nem folyamatosak a genomban, hanem úgynevezett intronok megszakítják meg ezeket. A gének tagolódása exonokra és intronokra egy sor evolúciós kérdést is felvetett. Felvetődik a kérdés, hogy vajon "Miért vannak a gének darabokban?" (56). A lehetséges válaszok között szerepelt, hogy a splicing, a pre-mRNS érési folyamata egyedülálló, és főképpen nagyon gyors evolúcióra ad lehetőséget. Az egyszeri pontmutáció, amely eddig maximum egy aminosav kicserélődését okozhatta, az eukarióta evolúció során, ha például egy kivágódási helyet érint, egyből megváltoztathatja a végtermék fehérjét a kódoló régió kihasadásával, vagy éppen egy új exon beékelődésével. Továbbá, a kivágódási helyet ért mutáció nem biztos, hogy száz százalékos hatású, így a szervezetnek egyszerre többféle termék állhat rendelkezésére kipróbálásra. Az intronok által megnövelt génekben nagyobb az esély a rekombinációs eseményekre, főleg, ha az intronokban különböző gyakoriságú és hosszúságú repetitiv szekvenciák is vannak.

Ezekkel a gondolatokkal együtt az is felmerült, hogy a géneknek az exonok különböző duplikálódásaiból, rekombinációiból történő felépülése, az exon shuffling, nem jelenti-e azt, hogy az exonok kódolják a fehérjék szerkezeti, vagy éppen funkcionális egységeit (57).

Az intronok létét magyarázandó, kétféle elmélet hódított tért. Az egyik szerint az intronok ősi gének maradványai, melyeket a prokarióták elvesztettek, míg az "önző" gén elmélet szerint az intronok, esetleg mint transzpozonok, a törzsfejlődés során ékelődtek az exonok közé, és az intronokban lévő DNS mennyisége ma is "önző" módon egyre növekszik (58).

A nyolcvanas években aztán kiderült, hogy valóban vannak olyan fehérjék, mint például a vérplazma proteázok vagy az immunglobulinok és általában az extracelluláris proteinek, amelyeknek a szerkezeti és/vagy funkcionális egységeit önálló exonok vagy exonpárok alkotják. Ezeknek az úgynevezett mozaik fehérjéknek a léte az első elméletet

erősíti, mármint azt, hogy az exonok már eredetileg is önálló funkcionális egységek voltak. Ugyanakkor a sejtben található fehérjék nagy részénél nem sikerült összefüggést találni az exonok és a funkcionális domének között, vagy egyáltalán nem is határozhatók meg a fehérjékben funkcionális domének. Ezek a példák a második, az "önző" DNS elméletet igazolják, miszerint az intronok később ékelődtek be a fehérjéket kódoló génekbe.

2. Mozaik fehérjék, exon shuffling

A mozaik fehérjékre jellemző, hogy modulokból épülnek fel, amely modulok más fehérjékben más modulokkal kombinálódnak. A modul a mozaik felépítésű fehérjék önálló feltekerődésre képes eleme. A moduláris, vagy mozaik felépítésű fehérjék génjeinek pedig közös jellemzője, hogy bennük az egyes modulokat független exonok, vagy exon csoportok kódolják. Ezeknek a fehérjéknek a kialakulása valószínüleg az exon shuffling során történt. Mi is ez a jelenség? Patthy László (59) írta le először 1987-ben az exon shuffling alapjait konkrét modellekkel igazolva. Modellként a vérplazma proteázait használta, amelyekről már korábban kimutatta, hogy domén felépítésüek (60). Az exon shuffling során három esemény történhet; az exonok beékelődése, kivágódása, vagy duplikációja. Ezekkel az eseményekkel lehet magyarázni, hogy a plazma proteázokban különféle modulok vannak, amelyek nemcsak funkcionális, hanem egy-egy globuláris strukturális egységet is alkotnak. Nagyon fontos az exon shuffling során, hogy egy új exon, amely beékelődik egy gén intronjába, a két határoló exonnal azonos fázisban szakítsa meg az intront (ez a feltétel értelemszerűen teljesül egy exon duplikáció során). Mit jelent ez, és miért fontos?

Az intronokat három csoportba lehet osztani aszerint, hogy a leolvasási kerethez képest hogyan helyezkednek el. Amennyiben az exont egy triplet első és második nukleotidja között szakítja meg az intron, akkor 1-es fázisúnak hívjuk; ha a második és a harmadik nukleotid között, akkor 2-es fázisúnak; ha két triplet között, akkor 0 fázisúnak. Értelemszerűen ha egy egyes fázisú intronba beékelődik egy exon, amelyik egy kettes fázisú intronból származik, akkor ez a fehérje szintézisekor elcsúszást fog okozni a leolvasási keretben. Az exon shuffling szabályaiból következik, hogy az azonos fázisú modulokat egy-egy csoportba lehet sorolni, mivel ezek képesek egymással kombinálódva új mozaik fehérjéket kialakítani (59).

Az exon shuffling evolúciós lehetőségeinek feltárásával együtt felvetődött, hogy tulajdonképpen korlátozott számú "ős" exon elégséges lehet-e az eukariótákban fellelhető sokféle fehérje kialakításához? Dorit és mti. (58) probálkoztak meg ennek a kérdésnek a megválaszolásával. Egy adatbázist készítettek az EMBL és a GenBank DNS

szekvenciabankokból, melyben szerepelt az addig ismert összes eukarióta exon. Az adatbázisban nem vették figyelembe a homológ gének, és a különböző fajok azonos génjeinek exonjait, azok csak egy példányban szerepelhettek. Ezután azt nézték, hogy egy exon milyen valószínűséggel fordul elő egy nem homológ másik génben, különböző összehasonlító programokat futtatva. A kapott adatokból aztán levezették, hogy szerintük az exon univerzum 1000-7000 darab exonból állhat.

Más kutatók arra tettek kísérletet, hogy összegyüjtsék azokat az eddig ismert fehérje modulokat, amelyek a különböző mozaik fehérjéket alkotják. Bork (61) 47 ilyen modult talált a különféle extracelluláris fehérjékben és receptormolekulákban, de ez a szám napról napra növekedik; ahogy újabb és újabb mozaik fehérjéket ismerünk meg.

3. A homológ modulok vizsgálatának a jelentősége és módjai

A homológ modulok összehasonlítása révén három fő következtetést vonhatunk le:

1. több homológ modul összehasonlításával olyan modul konszenzus, vagy profile állítható fel, aminek segítségével könnyebb megtalálni egy adatbankból a modul kisebb homológiájú rokonait. Hogyan készül a konszenzus? A homológ modulokat egymás alá írjuk és azt nézzük, hogy egy bizonyos pozicióban egy aminosav, vagy hasonló aminosavak hányszor fordulnak elő. Megállapítunk egy küszöb értéket, mondjuk 50%-ot, és az ezen érték fölötti gyakorisággal szereplő aminosavakat írjuk be a konszenzus szekvenciába. Ha egy új modult találunk, akkor azt is hozzá lehet írni a szekvenciákhoz és így tovább lehet finomítani a konszenzust is.
2. A több homológ modul egymáshoz igazítása, az úgynevezett multiple alignment során az előbb ismertetett módszerrel felírt konszenzus szekvenciák megmutatják, hogy melyek a modul legjobban konzerválódott részei. Ezek a konzervatív régiók, aminosavak kifejezhetnek egyrészt strukturális konzerválódást (mint például általában a ciszteinek előfordulása egy modulban), másrészt funkcionálisat is, rámutatva arra, hogy melyek azok az aminosavak, vagy régiók, amelyek a modul funkciójához esszenciálisak.
3. Mivel a homológ domének az exon shuffling elmélete szerint mind egy ős exonból duplikációk sorozatán keresztül keletkeztek, a modulok páronkénti összehasonlításai során nyert divergencia értékek arányosak azzal az idővel, amely eltelt a törzsfejlődésben az adott két modul duplikációja óta, ha a mutációk bekövetkezésének gyakoriságát viszonylag állandónak tekintjük. Ez a módszer hasonló ahhoz, ahogy például a globinok szekvenciájának összehasonlításával törzsfát

állítottak fel, csak itt egyszerre lehet horizontális (egyeden belüli), és vertikális (fajok közötti) változásokat is modellezni.

IV. A kapcsoló fehérjével homológ modulokat tartalmazó fehérjék

A kapcsoló fehérje tanulmányozása során hamar kiderült, hogy a C terminális részen lévő 200 aminosavnyi rész által képes kötődni a hialuronsavhoz (62), éppúgy, mint ahogy a proteoglükán is. Azonban kezdetben azt hitték, hogy ez a funkcióbeli hasonlóság nem párosul szekvenciahomológiával. Ezt támasztották alá egyes peptid térképezésből nyert adatok, és az is, hogy a kapcsoló fehérje hialuronsavkötő régiója ellen termelt ellenanyag nem adott keresztreakciót az aggrecannal. A csirke és a patkány kapcsoló fehérje gén cDNS szekvenciájának meghatározása azt mutatta, hogy a fehérje a C terminális részén egy kb. 100 aminosavból álló ismétlődést tartalmaz. A hialuronsavhoz a kapcsoló fehérje ezen ismétlődő doménjei kötődnek, és ezek a domének találhatóak meg az aggrecanban, és más hialuronsavhoz kötődni képes fehérjékben is. A gén exon-intron szerkezetének meghatározása és a fehérje másod- és harmadlagos szerkezetének vizsgálata nagyon egybevágó adatokat adtak; a kapcsoló fehérjében három hurkot lehet találni, melyeket három külön exon kódol, és az intronok mind egyes fázisúak (54).

1. Az első hurokkal rokon modulok

A kapcsoló fehérjét alkotó három önálló hurok (4. ábra) közül az első hurok szekvenciájáról hamar kiderült, hogy az homológiát, igaz nagyon kicsit, mutat az immunglobulinok variabilis régióiban, és a T sejt receptorokban található ugynevezett Ig-modulokkal (63). Ez a modul szintén megtalálható az aggrecanban (8, 64, 65), a versicanban (66), és a neurocanban (67). Újabban ugyanilyen Ig-modulokat találtak más, különböző adhéziós folyamatokban résztvevő fehérjékben is (68).

Mi lehet a szerepe az Ig-modulnak a kapcsoló fehérjében és a proteoglükánban? A vizsgálatok szerint az Ig-modul általában 90-100 aminosavból áll. Harmadlagos szerkezete viszonylag egyszerű, két antiparallel béta szál helyezkedik el egymás fölött szendvicsszerűen, és ezt egy diszulfid híd stabilizálja. Az immunglobulinok esetében ezek a modulok az antigénfelismerésben játszanak szerepet. A különböző adhéziós molekulákban, ahol szintén előfordulnak, fehérje-fehérje kölcsönhatásokat közvetítenek (68). A porcban szintén fehérje-fehérje kölcsönhatásban vesz részt a kapcsoló fehérjében és az aggrecanban található Ig modul azáltal, hogy ezen a modulon keresztül kötődik egymáshoz a kapcsoló fehérje és az aggrecan G1 egysége (52). Ennek a kötődésnek, aminek a természetéről még

nem sokat tudunk, nagy szerepe van a proteoglükán-hialuronsav aggregátum stabilizálásában, és így a porc rugalmasságának kialakításában.

2. A második és a harmadik hurokkal rokon modulok

Az aggrecanról készült elektronmikrószkópos felvételek, és más adatok is azt mutatták, hogy a hosszú fehérjegerincen három globuláris egység is található (2. ábra) (69, 70). Az N terminális felé eső első, a G1, teljes hosszában homológ a kapcsoló fehérjével, tehát tartalmaz egy Ig-modult, és két ismétlődő modult, amelyet egyesek PTR-nek neveznek (52). A G1 globuláris egységet egy rövid kar választja csak el a G2-től, amely szintén két példány PTR-t tartalmaz. Összesen tehát 4 darab PTR modul található egy aggrecan fehérjemolekulában. A G3-as egység lektinekkel homológ (19, 64 65). Ugyancsak mozaik fehérje a nagy fibroblaszt proteoglükán, a versican is (66). Ebben a fehérjében több egyéb modulon kívül szintén megtalálható egy Ig-modul és kettő PTR modul.

Az emberi agy és a marha gerincvelő fehérállományából izoláltak egy fehérjét, amely nagy affinitást mutat a hialuronsavhoz, ezért elnevezték GHAP-nak (71-73). A központi idegrendszer sejtközötti állományáról nagyon keveset tudunk, pedig nagyon fontos szerepe lehet például az idegrendszer regenerálódásában. Ezért is fontos az itt található fehérjék megismerése. Az innen izolált GHAP egyes darabjainak meghatározták az aminosavsorrendjét, és azt találták, hogy 100 %-osan megegyezik a versican Ig és PTR moduljainak megfelelő szakaszaival. Ebből arra következtetnek, hogy a GHAP tulajdonképpen a versicanból keletkezik proteolitikus hasítással (73). Erre azonban semmi közvetlen biznyíték sincs. Mindenesetre meglepő, hogy egy fehérje, amely az agy nedves tömegének egy ezrelékét teszi ki (tehát rendkívül gyakori), ugyanazokat a modulokat tartalmazza, mint a porcban található kapcsoló fehérje, vagy az aggrecan. A három fehérjében annyi a közös mai ismereteink szerint, hogy képesek a hialuronsavhoz kötődni. Az agy fehérállományában és a porcszövetben pedig az a közös, hogy ez a két hely a szervezetben, ahol az axonok nem képesek nőni. Felvetődik tehát a kérdés, hogy nincs-e ezeknek a fehérjéknek valami szerepük az axonok növekedésének gátlásában, annál is inkább, mivel kimutatták, hogy ehhez hialuronsavra van szükség (74)!

Egér agyban kimutatták, hogy legalább háromféle hialuronsavhoz kötődő fehérjének kell léteznie (75), és ezek közül az egyik kondroitin-szulfát oldalláncokat is tartalmaz. Valóban izoláltak is patkány agyból olyan cDNS klónokat, amelyek egy aggregálódó, kondroitin-szulfát oldalláncokat tartalmazó glükoproteint kódoltak. Ez a proteoglükán, a

neurocan, a versicannal mutat rokonságot, mivel ugyanazokat a modulokat tartalmazza, de nem azonos a GHAP fehérjével (67).

Nagy feltűnést keltett, amikor egyszerre két munkacsoport is közölte a cd44 antigén, vagy más néven limfocita "homing receptor" gén cDNS szekvenciáját (76, 77). Kiderült, hogy ez a fehérje az N-terminális extracelluláris részén szintén hordoz egy PTR modult, bár a homológia kisebb, mint például egy proteoglükán és egy kapcsoló fehérje PTR között. Érdekesség, hogy itt, az eddigi összes többi esettel ellentétben, csak egy példányban fordul elő ez a modul. Ugyancsak egy példányban fordul elő a TSG-6 fehérjében is, melynek szekvenciáját tumor nekrózis faktorral indukált sejtekből izolált cDNS-ből határozták meg (78).

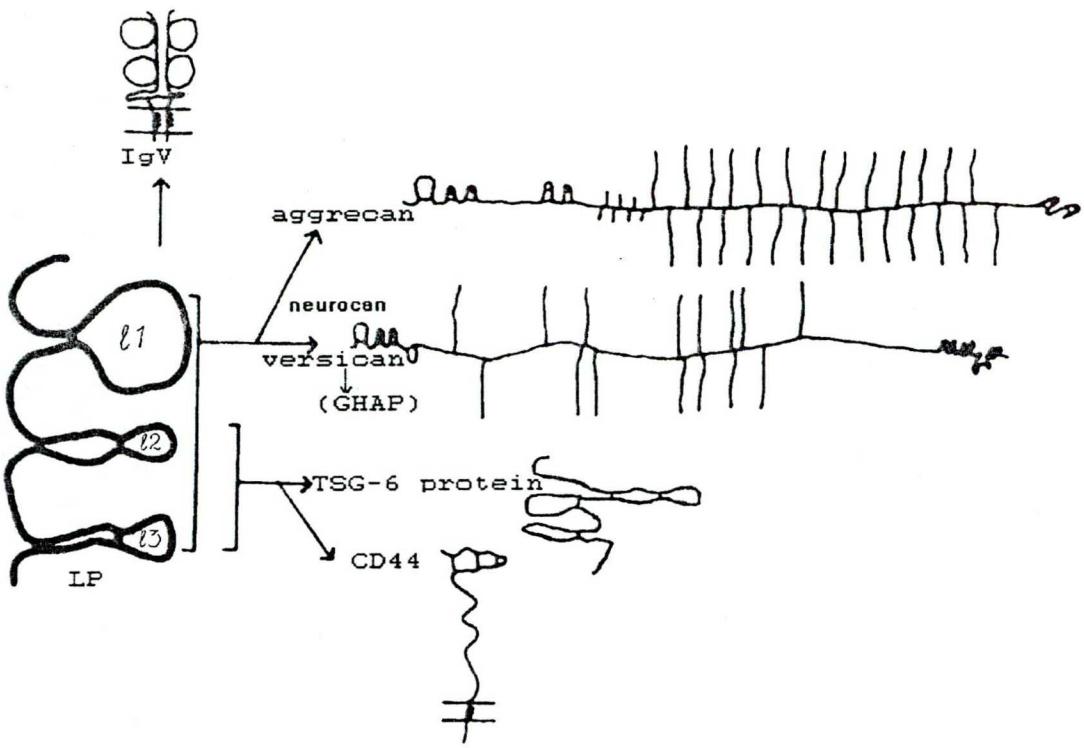
Felvetődik a kérdés, mi a közös az olyan, látszatra nagyon különböző rendeltetésű fehérjékben, mint a kapcsoló fehérje, az aggrecan, a versican, neurocan, a limfocita homing receptor, vagy éppen egy TNF indukált fehérje? Látszólag nagyon különböznek, azonban funkciójukat tekintve van egy közös tulajdonságuk, valami módon az összes felsorolt fehérje képes kötődni a hialuronsavhoz.

Az aggrecan G1 egysége, és a kapcsoló fehérje is a porcban kötődik a hialuronsavhoz, és ebben mindkét PTR modul részt vesz. Érdekes módon az aggrecan G2 egységében található két PTR modul az egyedüli kivétel, melyekről nem tudták kimutatni, hogy kötődik a hialuronsavhoz, valószínűleg csak térkitöltő szerepük van (52).

A limfocita homing receptor, vagy cd44 antigén a limfocitáknak a vénák falát borító endotélium sejtjeihez való adhézióját, és ezáltal a nyirokkeringésbe való kilépését segíti elő. A vizsgálatok szerint a hialuronsav ebben a folyamatban jelentős szerepet játszik, hiszen tulajdonképpen a cd44 a hialuronsav specifikus receptora (79). Kimutatták, hogy a tumoros fehérvérsejtek közelében megnövekszik a hialuronsav koncentrációja, és ilyenkor a sejtek felszínén a limfocita homing receptor bizonyos módosult formái jelennek meg (80). Tehát van valami összefüggés a sejtek daganatos elváltozása, és a hialuronsav között. Ezért is gondolják, hogy a TNF által indukált sejtekből kiszabaduló TSG-6 fehérje, ami szintén nagy affinitással kötődik a hialuronsavhoz, egyféle kompetítora lehet a módosult cd44 formáknak, és így közrejátszhat a rákos elfajulás megszüntetésében (78).

Tehát megállapíthatjuk, hogy a szervezetben a hialuronsavnak fontos szerepe van, és ahol ez a molekula megtalálható, ott minden bizonnyal fellelhetőek a hialuronsavat kötő fehérjék is. Ezekben a fehérjékben pedig az a közös, hogy valamennyien az exon shuffling során nyerték el mai szerkezetüket, és valamennyiben megtalálható egy vagy több PTR modul, amely a kötődésért felelős.

A kapcsoló fehérjével rokon modulokat tartalmazó fehérjék szerkezetét összegzi a 6. ábra.



6. ábra: A kapcsoló fehérjében található modulokkal rokon modulokat tartalmazó fehérjék. IgV: immunglobulinok variabilis lánc, LP: kapcsoló fehérje.

CÉLKITŰZÉS

A gerincesek porcszövete rendkívül rugalmas, ugyanakkor nyomással szemben ellenálló képződmény. Ezen a tulajdonságait a sejtközötti állomány megfelelő összetétele határozza meg. Az itt található extracelluláris fehérjék közül a kapcsoló fehérje részt vesz a hialuronsav aggregátumok képzésében azáltal, hogy kötődik mind a hialuronsavhoz, mind a nagy aggregálódó, kondroitin-szulfát oldalláncú proteoglükánhoz, az aggrecanhoz. A Kötőszövet csoport tagjai izolálták a kapcsoló fehérje kódoló részét, és a 3' nem kódoló régiót átfedő cDNS klónokat, valamint genomiális klónokat. Az 5' végen azonban az egyik cDNS klón olyan DNS darabot tartalmazott, amely nem volt reprezentálva az izolált genomiális klónokban. Ez alapján feltételezhető volt, hogy létezik még egy vagy több exon az 5' végen. A csoport további munkájához ezenkívül feltétlenül szükséges volt a kapcsoló fehérje gén 5' végi feltételezett szabályozó régióinak az izolálása. Ennek érdekében határoztuk el, hogy azonosítjuk a kapcsoló fehérje gén 5' végét, izoláljuk a legelső exont vagy exonokat reprezentáló cDNS klónokat, valamint az ezeket, és a promoter régiót tartalmazó génszakaszt, továbbá meghatározzuk a transzkripció kezdőpontját.

Ezen célkitűzéseket az alábbi kísérletek elvégzésével kívántuk megvalósítani:

1. Specifikus oligonukleotid primerek segítségével klónozzuk az 5' végnek megfelelő cDNS darabokat, meghatározzuk ezek szekvenciáját.
2. Az cDNS klónokat felhasználva izoláljuk az első exont vagy exonokat, valamint a promoter régiót hordozó genomiális DNS darabot, és jellemezzük azt.
3. Meghatározzuk a transzkripció kezdőpontját, vagy kezdőpontjait.
4. Megvizsgáljuk, hogy csirkében is létezik-e a patkány kondroszarkoma kapcsoló fehérjében leírt alternatív exon.

A kapcsoló fehérje egy mozaik fehérje, három önálló modulból áll, amelyeket önálló exonok kódolnak. Rokon modulok találhatóak különféle egyéb mozaik fehérjékben is. A rokon modulok összehasonlítása értékes adatokkal szolgáltathat az exon shuffling folyamatáról. Ezért határoztuk el, hogy a különböző rokon modulok számítógépes vizsgálata alapján egy modellt állítunk fel, mellyel megpróbálunk magyarázatot adni arra a felvetésre, hogy milyen sorrendben kezdődhetett el az egyes, rokon modulokat hordozó fehérjék evolúciója.

ANYAGOK ÉS MÓDSZEREK

I. Kísérleti anyagok, baktérium törzsek, plazmidok, fágok

1. Vegyszerek

A kísérletekhez használt finomvegyszerek Sigma, Merck, Serva, vagy Reanal gyártmányúak voltak. A restrikciós endonukleázokat és az egyéb enzimek nagy részét vagy a New England Biolabs-tól vettük, vagy az Szegedi Biológiai Központ Biokémiai Intézetének munkatársai tisztították és bocsátották rendelkezésünkre. A reverz transzkripcióhoz használt Mo-MuMLV reverz transzkriptázt a BRL-től vettük. A T4-fág polinukleotid kináz és a T4-fág polimeráz, a nem rádioaktiv dezoxi- és dideoxiribonukleozid-trifoszfátok, a Sephadex G-50 és a Sepharose CL-4B kromatográfiás gélek a Pharmacia-PL Biochemicals gyártmányai voltak. Az [α-^{32}P]-dNTP-t az MTA Izotóp Intézetének Izinta Izotópkereskedelmi Leányvállalata szállította. A [γ-^{32}P]-dATP-t az SzBK-ban állították elő. A hibridizációhoz használt nitrocellulóz filterek Schleicher & Schuell, valamint Millipore gyártmányúak voltak. Az oligonukleotidokat az SzBK Biokémiai Intézetében Szilák László szintetizálta foszforamidit módszerrel (81). A dolgozatban felhasznált primerek láthatóak az 1. táblázatban.

1. táblázat: a szintetikus oligonukleotid primerek szekvenciái
A számozás a 14. ábra szerint történt. A PEr primer a pLP8G1 klón szekvenciáját követi.

név	szekvencia	pozíció
	Normál primerek	
PA	5′-CAAGTGGTCAGAAGTATCACGT-3′	-605 - -584
rpr404	5′-TAGTTCGGGACTGGTGTGCG-3′	-488 - -469
rpr846	5′-ACTTGGGAGCTCCACACAA-3′	-46 - -28
	Fordított primerek	
PEr	5′-GGTGGCAGAGGAC.CTAAAAAAGCTCTGCAC-3′	-176 - -188
	és	-452 - -468
PFr	5′-CTGCCCCAGCCTCCGTGCCCCTCAGCTTCT-3′	-404 - -433
prE2	5′-CTTGTCATCTTCACAGTCAC-3′	8 - -12
prE3	5′-ACCACAAGTAGGCGGGGTCC-3′	128 - 109

2. törzsek, vektorok

Baktérium törzsek:

Q359: *recA*$^+$, *hsdR*$_k^-$, *hsdM*$_k^+$ *supE*, $\phi 80^r$ (82)

DH-1: *F*$^-$, *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, (r_k^-, m_k^+), *supE44*, *relA1* (82)

DH-5 α: *F*$^-$, *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, (r_k^-, m_k^+), *supE44*, *relA1*, Δ(*lacZYA-ArgF*) *U169* [ϕ*80 dlac* Δ(*lacZ*)*M15*] (83)

JM101: *supE*, *thi*, Δ(*lac-proAB*) [*F' TraD36*, *ProAB*, *lacI*^{q}*Z*Δ*M15*] (82)

JM107: *endA1*, *gyrA69*, *thi*, *hsdR17*, *supE44*, *relA1*, Δ(*lac-proAB*) [*F' TraD36*, *ProAB*, *lacI*^{q}*Z*Δ*M15*] (82)

vektorok: pUC8, pUC9, pUC12 (84); M13mp18, M13mp19 (85)

3. Táptalajok

LB: 10 g Bacto triptont, 5 g Bacto élesztő kivonatot és 10g NaCl-ot tartalmaz literenként, (pH 7.2).

SOB: 20 g triptont, 5 g élesztőkivonatot, 2 ml 5M NaCl-ot és 2.5 ml 1M KCl-ot tartalmaz literenként.

2YT: 16 g triptont, 10 g élesztőkivonatot, 5 g NaCl-ot tartalmaz literenként.

LA alapagar: LB kiegészítve 1.25-2% Bacto agarral

LA fedőagar: LA alapagar kiegészítve 0.6% Bacto agarral

Az ampicillint 100μg/ml végkoncentrációban alkalmaztuk.

4. Oldatok

TEN: 40 mM TrisHCl, 1 mM EDTA, 0.15 M NaCl (pH 7.5).

1x MOPS: 50 mM MOPS, 1 mM EDTA, (pH 7.0)

1x SSPE: 0.15 M NaCl, 10 mM $NaH_2PO_4{\cdot}H_2O$, 1 mM EDTA, (pH 7.4).

1x E puffer: 50 mM Tris, 2 mM EDTA, 10 mM Na-acetát (pH 8.05).

1x SSC: 0.15 M NaCl, 0.015 M trinátrium citrát (pH 7.0).

NET: 0.15 M NaCl, 0.1 mM EDTA, 20 mM TrisHCl (pH 8.0).

100x Denhardt oldat: 2% Ficoll, 2% Polivinil-pirrolidon, 2% BSA.

TFB : 10 mM MES·KOH, 100 mM RbCl, 10mM $CaCl_2{\cdot}H_2O$, 45 mM $MnCl_2{\cdot}4H_2O$, 3 mM $[NH_2]_6CoCl_3$, (pH 6.15).

STET: 50 mM Tris, 50 mM EDTA, 8% szacharóz, 5% Triton X-100, pH=8.0

SM: 0.1 M NaCl, 10 mM $MgSO_4$ 50 mM TrisHCl, 0.01% zselatin, pH=7.5

6xSET 0.9 M NaCl, 0.18 M Tris, 6 mM EDTA, pH=8.0

II. MÓDSZEREK

1. DNS izolálás

A plazmid DNS-eket LB ampicillin táptalajban éjszakán át történő intenzív rázatással felnövesztett sejtekből preparáltuk.

"Plazmid miniprep" (86): 1,5 ml kultúrát Eppendorf-centrifugacsőben lecentrifugáltunk. A sejteket 200 µl STET pufferben 15 µl friss 10mg/ml lizozim jelenlétében tártuk fel, majd a sejttörmeléket centrifugálás után fogpiszkálóval távolítottuk el. A plazmid DNS-t izopropanollal kicsaptuk, majd centrifugálás, mosás, szárítás után 25µl 50 µg/ml RN-ázt tartalmazó TE pufferben oldottuk fel.

Nagy mennyiségű plazmid DNS kinyerése céljából 1l bakteriumkultúrában a plazmidokat kloramfenikollal amplifikáltuk, majd Clewel és Helinski módszerével tártuk fel a sejteket (87). A törmelékmentes sejtlízátumot fenol és kloroform 1:1 arányú elegyével extraháltuk, RN-ázzal kezeltük, majd kicsapás után a plazmid DNS-eket CsCl-etidiumbromid egyensúlyi sűrűség gradiens ultracentrifugálással tisztítottuk tovább (88).

2. M13 fág DNS izolálás

A szekvenálni kívánt fragmentumokat M13 fág vektorokba szubklónoztuk Maniatis és mti. útmutatásai alapján (82). A klónozáshoz az M13mp18 és M13mp19 fágvektorok replikatív formáit kellett izolálni. Ugyancsak replikatív formát kellett izolálni a rekombináns fágok restrikciós emésztéssel történő ellenőrzéséhez. A szekvenáláshoz, T4 DNS polimerázzal történő térképezéshez pedig egyszálú fág DNS-t kellett kinyernünk.

A ligálás során nyert rekombináns fág DNS-t Hanahan (83) módszerével juttatuk be JM107 baktériumsejtekbe, majd IPTG-t és X-gal-t tartalmazó fedőagarral lemezeltük ki. A szintelen plakkokból nyert fágokkal 50 µl JM107 kultúrát fertőztünk 2ml LB tápoldatban. 4-5 órás 37 °C-on történő rázatás után a kulturából kiülepítettük a bakterium sejteket és a felülúszót tekintettük a továbbiakban fágtörzsnek. A szuszpenzió titerét hígítási sor kilemezelésével határoztuk meg.

Egyfonalas fág DNS tisztításához a fágrészecskéket PEG 8000 oldattal kicsaptuk NaCl jelenlétében. A fágfehérjéket többszöri fenolos extrakcióval távolítottuk el. A DNS-t etanollal kicsaptuk, szárítottuk, majd TE pufferben szuszpendáltuk.

Replikatív forma izolálásához 5, illetve 500 ml térfogatban 2YT tápoldatban felnövesztett JM101 sejteket fertőztünk $1x10^{10}$-$5x10^{12}$ pfu/ml titerű fágszuszpenzió 0.2, illetve 2 ml-ével. 15 percig rázattuk a tenyészeteket, majd kloramfenikol antibiotikum hozzáadásával gátoltuk az egyszálú fágforma képződését, és további 1-2 órát rázattuk. A fág

replikatív formáját tartalmazó baktériumsejteket kiülepítettük, majd a továbbiakban az alkali plazmidtisztítási módszer szerint izoláltuk a fág replikatív forma DNS-t (82).

3. lambda fág DNS izolálás (82)

A lambda fágokat Q359 baktérium gazdasejteken szaporítottuk. A plakkokból leszívott fágrészecskékkel, éjszakán át maltóz és Mg^{++} jelenlétében növesztett, fél térfogat 10 mM $MgSO_4$ oldatban felszuszpendált sejteket fertőztünk, majd LA fedőagarral kilemezeltük. A fágot olyan koncentrációban alkalmaztuk, hogy a lemezen a tarfoltok majdnem összeérjenek. A plakkokból a fágokat SM pufferrel összemostuk, és meghatároztuk a szuszpenzió titerét. Nagyobb mennyiségű fág izolálásához Q359 sejteket 10 mM Mg^{++} tartalmú LB táptalajon OD_{650}=0.45-0.6 értékig növesztettünk, és úgy fertőztük a fágokkal, hogy megközelítőleg egy baktériumsejtre egy fág jusson. A fertőzött sejteket addig rázattuk, amíg a teljes lízis be nem következett. Ezt az oldat feltisztulása jelezte, amit a 650 nm-en mért fényelnyelés csökkenése kisért. A kultúrából kiülepítettük a sejttörmeléket (5000 rpm, 10 perc, 4 °C), a fágrészecskéken kívüli nukleinsavakat DN-áz és RN-áz enzimekkel emésztettük. A fágokat NaCl jelenlétében 10% PEG-gel kicsaptuk és lecentrifugáltuk (7000 rpm, 15 perc, 4 °C). A csapadékot SM pufferben szuszpendáltuk fel, majd a PEG-et kloroformos extrakcióval távolítottuk el. A fágokat CsCl lépcsős grádiensen tisztítottuk tovább (SW27 rotor, 20000rpm, 2 óra 10°C). Az 1.5 g/ml sűrűségű rétegben öszegyűlt fágokat fecskendővel leszívtuk. A CsCl-ot 0.1% SDS-sel kiegészített DNS dializáló pufferrel (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA) szemben dializáltuk többszöri puffercserével. A fágfehérjéket proteináz K-val emésztettük (37°C, 1 óra), majd fenolos extrakcióval eltávolítottuk. A fenol nyomoktól újabb több napos dialízissel szabadultunk meg.

4. Restrikciós emésztés

Plazmid és fág DNS analítikai emésztésénél általában 1 µg DNS-t és 1-3 egység enzimet használtunk 20 µl végtérfogatban, az enzim előállítójának tanácsai szerint. Preparatív célú emésztéseknél 5-50 µg DNS-t használtunk nagyobb térfogatban. Kettős emésztéseknél, ha a két puffer eltért egymástól, az első emésztés után kicsaptuk a DNS-t és a feloldás után az új pufferben emésztettük tovább.

5. DNS fragment izolálás

A preparativ emésztés után a kiválasztott DNS fragmentum elé egy kis darab Schleicher & Schuell NA-45 DEAE membránt ékeltünk a gélbe, és a fragmentumot erre

futtattuk rá. A membránról az ionosan kötött DNS-t 65 °C-on történő 1.5 M-os NaCl-t tartalmazó NET pufferben oldottuk le, majd alkohollal csaptuk ki egy fenol kloroformos extrakció után.

6. Rádioaktív próbák jelölése

A hibridizációs kísérletekhez a próba jelölését Feinberg és Vogestein (89) módszere szerint random primer meghosszabbításával végeztük. 20-100 ng fragmentumhoz denaturálás után 5- 10-szeres moláris feleslegű random hexanukleotid primert hibridizáltunk és 50-70μCi [α-^{32}P]-dATP jelenlétében *Escherichia coli* DNS polimeráz Klenow fragmentumával hosszabítottunk meg. A be nem épült nukleotidokat a próbától Sephadex G50 oszlopon történő gélszűréssel távolítottuk el.

Az oligonukleotidokat [γ-^{32}P]-dATP-vel jelöltük polinukleotid kinázzal Maxam és Gilbert módszerével (90).

7. DNS elektroforézis

Az 500 bp-nál nagyobb DNS fragmentumokat 0.6-1.5%-os horizontális agaróz gélen választottuk el. Molekulasúly markerként *Hin*dIII enzimmel hasított vad típusú lambda fág DNS-t használtunk. Az 500 bp-nál rövidebb DNS fragmentumokat 5%-os vertikális poliakrilamid gélen választottuk el egymástól, *Bsp*I enzimmel emésztett pUC12 plazmid DNS molekulasúly marker mellett. A DNS-t 0.5μg/ml ethidium-bromiddal festettük és 254 nm-es ultraibolya fénnyel tettük láthatóvá.

8. RNS izolálás

Össz RNS izolálásához 14 napos csirke embriók mellcsont kezdeményeit megtisztítottuk a perikondriumtól, majd folyékony levegőben lefagyasztottuk és -70°C-on tároltuk a felhasználásig. A szövetet politron típusú homogenizátorral tártuk fel Adams és mti. (91) módszere szerint guanidium-hidroklorid jelenlétében. Az RNS-t három egymást követő etanolos kicsapással választottuk el a DNS-től és a fehérjék zömétől, majd kloroform:izobutanol 4:1 arányú keverékével zsírmentesítettük és fenollal extraháltuk. Kétszeri oligo(dT)-cellulóz kromatográfiával (92) poly(A)$^+$ RNS frakciót állítottunk elő, amit a kicsapás és feloldás után használtunk a további kísérletekhez. A koponya tető poly(A)$^+$ RNS W.S. Argraves ajándéka volt.

9. Northern hibridizáció

RNS elektroforézishez 1-2 µg poly(A)$^+$, vagy 10-15 µg össz RNS-t használtunk. Felvitel előtt az RNS-t 50% formamid jelenlétében denaturáltuk, majd 0,8% agaróz, 2,2M formaldehid, 1xMOPS tartalmú géleken választottuk el, a gélt 10xSSC-ben rázattuk 30 percig, és az RNS-t kapilláris úton transzferáltuk nitrocellulóz filterre. A hibridizációt 50% formamid, 2xSSC, 0.1%SDS, 0.1mg/ml hering spermium DNS tartalmú elegyben végeztük 43 °C-on éjszakán át, 12 órás előhibridizáció után. A filtereket 1xSSC, 0.1%SDS, 0.05% nátrium-pirofoszfát oldattal mostuk 68°C-on 4-szer 15 percig, majd megszárítottuk és autoradiografáltuk.

10. Primer extenzió

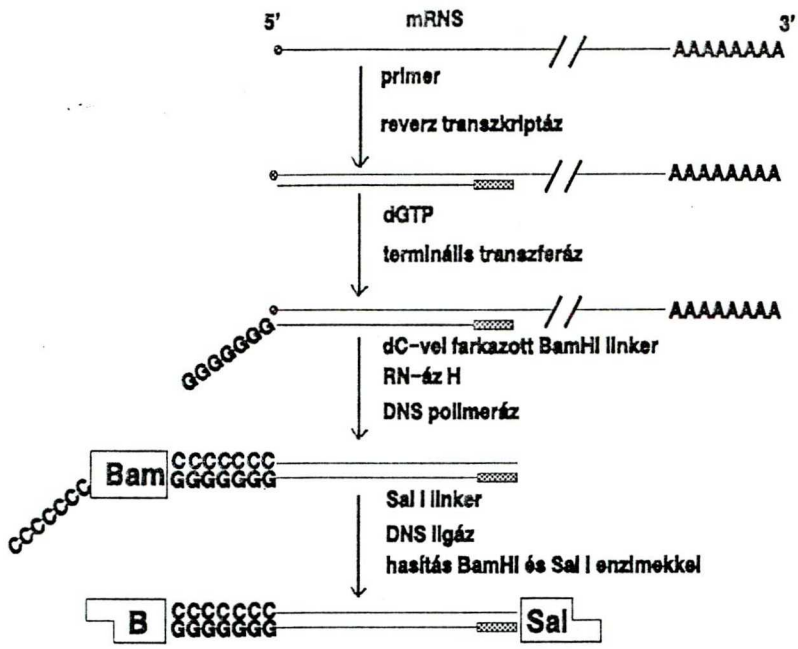
A primer extenzióhoz 0.1 pmol végjelölt szintetikus oligonukleotid primert hibridizáltunk 4 µg 10 mM metil-merkuri-hidroxiddal denaturált poly(A)$^+$ RNS-sel, majd a primereket 40 egység Mo-MuMLV reverz transzkriptázzal hosszabbítottuk meg 50 µl reakcióelegyben (50 mM TRIS-HCl (pH=8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.5 mM dNTP, 40 µg/ml actinomycin D, 20 egység ribonukleáz inhibitor) 42 °C-on 40 percig. A mintákat kicsapás után rádioaktiv molekulasúly marker mellett 5-7%-os denaturáló poliakrilamid gélen futtatuk, majd a gélt autoradiografáltuk.

11. cDNS klónozás

A cDNS klónozás menete látható a 7. ábrán. Az első lépés megegyezett a primer extenziónál ismertetettekkel, azzal a különbséggel, hogy 0.25 pmol prE2 és prE3 primert hibridizáltunk 5-5 µg poly(A)$^+$ RNS-sel. A komplementer DNS 3' végére oligo-dG farkat szintetizáltattunk terminális transzferázzal, majd oligo-dC farokkal ellátott *Bam*HI linkereket hibridizáltunk hozzá. A reakcióelegyhez RN-áz H enzimet adtunk, majd a második szálat 10 egység DNS polimeráz I-gyel szintetizáltattuk. A túlnyúló egyszálú végeket T4 DNS polimerázzal leemésztettük, illetve feltöltöttük. A duplaszálú cDNS-hez foszforilált *Sal*I linkert ligáltunk, majd *Bam*HI és *Sal*I enzimekkel emésztettük. A cDNS-t Sepharose CL4B gélen frakcionáltuk, hogy a nagyon rövid cDNS-ektől és a linkerektől megszabaduljunk, és *Bam*HI és *Sal*I enzimekkel hasított pUC12 vektorba ligáltuk.

12. Telephibridizáció

Telephibridizáció során Hanahan és Meselson (93) módszere szerint a kiszélesztett, vagy kiszurkált transzformánsokat hordozó lemezről nitrocellulóz filterrel replikát készítettünk, majd a filtert 0.5M NaOH oldattal kétszer, 1M TRIS-HCl pH=7.4, és 0.5M TRIS-



7. ábra: A cDNS klónozás menete. Bam, B: *Bam*HI linker, Sal: *Sal*I linker

HCL pH=7.4, 1.5M NaCl oldattal kezeltük. A filtereket 20 percig infralámpa alatt, majd 2 órán át 80°C-on vákuumban megszárítottuk. Ezt követően 65°C-on 2 óráig mostuk a filtereket 5x Denhardt, 0.5% SDS oldatban. Ezt követően 6x SET, 1x Denhardt, 0.1% SDS oldattal előhibridizáltuk, hozzáadtuk az előzőleg denaturált radioaktív próbát (rendszerint $2x10^6$ dpm-t), és egy éjszakán át hibridizáltuk 65°C-on. Másnap 2x SSC, 0.1% SDS oldattal mostuk 65°C-on négyszer 20 percig. Szárítás után autoradiografáltuk.

Szintetikus oligonukleotid próba estében a filtereket 6x SSC, 0.1%SDS, 1x Denhardt, 0.05% nátrium-pirofoszfát, 10% dextrán-szulfát elegyben előhibridizáltuk 37 °C-on 6 órán át, majd filterenként 2 ml fenti elegybe adtuk a végjelölt oligonukleotid próbát. 16-20 órai, 37 °C-on végzett hibridizálás után a filtereket négyszer 20 percig 37 °C-on és kétszer 30 percig 45 °C-on mostuk 6x SSC, 0.1% SDS, 0.05% nátrium-pirofoszfát oldatban.

13. Southern transzfer és hibridizáció (82)

Agaróz gélelektroforézist követően a DNS sávokat nitrocellulóz filterre transzferáltuk. 5 kb-nál nagyobb fragmentumok transzferének megkönnyítése érdekében a gél megfelelő

részét 0.25 M HCl oldatba áztattuk 20 percig, majd desztillált vízzel leöblítettük. A DNS kettős szálát denaturáló oldattal (1M NaCl, 0.5 M NaOH) választottuk el egymástól 2x20 perces áztatással. Ezt követően a gélt semlegesítő oldatban (3 M NaCl, 0.5 M Tris-HCl pH 7.0) áztattuk, majd a DNS-t a kapillaritás elve alapján nitrocellulóz filterre transzferáltuk. Poliakrilamid gélek esetén a gélt 10 percig forraltuk 1M ammónium-acetát, 20mM Natrium-hidroxid elegyben, majd az agaróz gélekhez hasonlóan transzferáltuk a továbbiakban. Szárítás után a DNS-t 80 °C-on vákuumban 2 órán keresztül sütöttük rá a filterre.
A filtert 5xSSC, 5xDenhardt, 100 µg/ml hering sperma DNS elegyében 68°C-on 2-4 órán keresztül előhibridizáltuk. Utána hozzáadtuk a radioaktívan jelölt denaturált próbát, és további 16 órán keresztül inkubáltuk. Ezután a filtert 2xSSC-ben szobahőmérsékleten, majd 2xSSC-ben 4x15 percig, 1xSSC-ben 2x15 percig, 0.2xSSC-ben 15 percig 65°C-on mostuk és infralámpa alatt megszárítottuk. Az autoradiográfiát erősítő ernyő segítségével Forte Medifort röntgenfilmmel végeztük -70°C-on.

14. Szekvenálás

A cDNS klónok szekvenciáját Sanger és mtı. (94) szintézis terminációs módszerével határoztuk meg. A klónok inszertjeit M13mp18 és M13mp19 fágvektorokba klónoztuk. Primerként az univerzális, valamint a prE3, és prE2 primereket használtuk. A reakció jelöléséhez [α-^{32}P]-dATP-t használtuk, a mintákat 50 watt teljesítménnyel futtatuk 6, illetve 8%-os denaturáló poliakrilamid gélen, majd beszárítás után a gélt autoradiografáltuk. A pLPD1 klón szekvenciáját Deák Ferenc és Markus Biesold határozták meg.

15. Polimeráz láncreakció (PCR)

A PCR-ben templátként a primer extenzió termékét, vagy kontrollként a megfelelő plazmid DNS 1 ng-ját használtuk 100 µl végtérfogatban. A reakció elegyben 100-100 pmol megfelelő primer, 0.2 mM dNTP, 67 mM TRIS-HCl (pH=8.8), 16.6 mM $(NH_4)_2SO_4$, 10 mM 2-merkapto-etanol, 6.7 mM $MgCl_2$, 10% (v/v) DMSO, és 2.5 egység *Taq* DNS polimeráz volt. A párolgás megakadályozására az elegyet paraffin olajjal fedtük be, amit a végén kloroformos extrakcióval távolítottunk el. Az 5 perces 94 °C-on történő denaturálás után 30 ciklus következett (2.5 perc 48 °C-on, 3.5 perc 74 °C-on és 1.5 perc 94 °C-on), majd a végén egy 10 perces elongáció 74 °C-on. Amennyiben szükséges volt, a PCR eredményének egy századát templátként felhasználva újabb PCR-t végeztünk, vagy az utolsó 5 ciklusban végjelölt primereket adtunk a reakcióhoz, hogy autoradiográfiával az alacsony koncentrációjú termékeket is ki tudjuk mutatni.

16. T4 DNS polimeráz térképezés

A mRNS-hez képest komplementer szálat tartalmazó egyszálú M13 DNS-hez mRNS-t, és végjelölt PA primert hibridizáltunk, majd a primert T4 DNS polimerázzal meghosszabbítottuk Hu és Davidson módszere alapján (95). A termékeket denaturáló poliakrilamid gélen futtattuk szekvenciamarker jelenlétében.

17. Komputeres analízis

A számítógépes analízist egyrészt az IBM PC kompatibilis gépeken futó, az Egyesült Államokból, a Cornell egyetemről származó programokkal, másrészt a Gödöllői Mezőgazdasági Biotechnológiai Kutatóközpontban rendelkezésünkre álló, MicroVAX számítógépen futó GCG programcsomaggal (96) végeztük. DNS illetve fehérje adatbázisként az EMBL, a GenBank, illetve a SwissProt és a PIR adatbázisokat használtuk, amelyek Gödöllőn álltak a rendelkezésünkre.

A programok közül ki kell emelni a Fasta és a ProfileSearch programokat, melyekkel az adatbázisokban kerestünk, a Gap és a PileUp programokat melyekkel az összehasonlításokat végeztük, valamint a Fold programot, mellyel a másodlagos szerkezet predikciókat végeztük Zuker algoritmusa alapján (97).

EREDMÉNYEK

I. cDNS klónozás

Az irodalomban kevés példát lehet találni a mRNS 5' végét képviselő cDNS irányított klónozására, mert általában az oligo dT-vel, vagy random primerrel indított cDNS klónozási módszerekkel elegendő hosszúságú cDNS-eket lehet kapni a további munkához, így a promoter régiót hordozó genomiális DNS klón megtalálásához, majd a transzkripciós kezdőpont meghatározásához. A csirke kapcsoló fehérje gén esetében a hosszú és repetitív DNS-t tartalmazó intronok megakadályozták, hogy átfedő genomi klónok kikeresésével jussunk el a gén 5' végéhez. A korábban vizsgált cDNS klónok az 5' nem transzlált régióból, és így az első exonból csak egy rövid szakaszt tartalmaztak. Ezért határoztuk el, hogy a megfelelő specifikus primerek alkalmazásával célzottan az első exont klónozzuk.

1. A mRNS tisztítása

A mRNS-t 14-17 napos csirke embriók mellcsont kezdeményéből tisztítottuk. Ilyen korban a csirkeembriók mellcsontjai még teljesen porcosak és kifejeződnek bennük a porcspecifikus gének, köztük a kapcsoló fehérje génje is. A mellcsont kezdeményeket teljesen megtisztítottuk az egyéb szövetdaraboktól, hogy csak porcspecifikus hírvívők legyenek a kinyert RNS-ben.

Az RNS tisztítását a szövetben lévő proteoglükán aggregátumok tették nehézzé. Ezért Adams és mti. (91) által speciálisan a porcszövetre kidolgozott RNS kinyerési módszert alkalmaztuk. A kinyert RNS mennyiségét fotometriásan (260 nm-en mért elnyelés), minőségét pedig agaróz gélelektroforézissel, és Northern hibridizációval ellenőriztük folyamatosan a tisztítás során. A Northern hibridizációhoz próbaként a Deák F. és mti. (46) által korábban izolált kapcsoló fehérjét kódoló cDNS klónok közül a pLPF4-et használtuk. A kapott össz RNS-ből kétszeri oligo-dT cellulóz kromatográfiával tisztítottuk a poly(A)$^+$ RNS-t. Végül is 50 µg poly(A)$^+$ RNS-t sikerült kinyerni.

2. A primer extenzió

Mivel célunk az volt, hogy a kapcsoló fehérje gén 5' végére specifikus cDNS klónkönyvtárat készítsünk, az 5' irányban leghosszabb cDNS klón, a pLPF4 szekvenciája alapján két olyan 20 tagú egyszálú oligonukleotidot szintetizáltattunk (1. tábl. prE3 és prE2), amelyek komplementerek az RNS-sel, és így hozzáhibridizálódva a hírvívő RNS-hez a reverz

transzkriptáz enzim számára primerként szolgálhatnak. A két primer közül az egyik, a prE2 az 5' véghez közelebb van, tehát ha erről indul az átírás, akkor rövidebb terméket kapunk, ami technikailag előnyösebb, mert így könnyebb megállapítani az átírt szekvencia hosszát. A másik primert, ami a 3. exon 5' végéhez hibridizál, azért használtuk, mert Rhodes és mtí. (55) patkány kondroszarkóma sejtekből izolált kapcsoló fehérje klónoknál egy alternatív exont mutattak ki a 2. és a 3. exon között. Ha ez az extra exon csirkében is létezik, akkor a prE3 primerről indított klónok között kell ilyet találnunk.

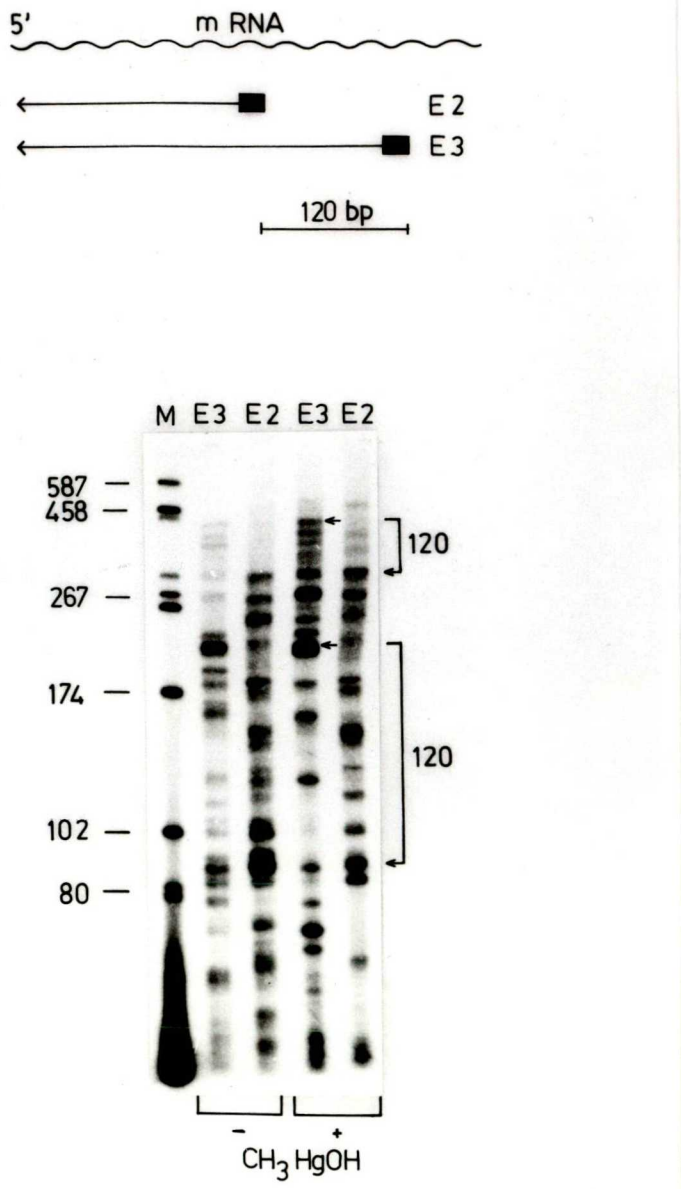
A cDNS klónozás első lépése a reverz transzkripció, amit a klónozástól függetlenül is elvégeztünk abból a célból, hogy megállapítsuk, milyen messze van a primerek hibridizációs helyétől az RNS 5' vége, azaz az átírás kezdőpontja. Mivel a két primer hibridizálási helye között 120 bp különbség van, ezért azt vártuk, hogy a prE3 primer meghosszabításával kapott DNS 120 nukleotiddal hosszabb lesz, mint a prE2 primeré. A 8. ábrán látható, hogy a primer extenzió során nem egy -az RNS 5'végéig átírt- termék keletkezett, hanem több. Ennek a következő okai lehetnek:

1. a primer nem csak a kapcsoló fehérje gén RNS-éhez hibridizált hozzá, hanem más RNS-ekhez is. Figyelemre méltó azonban, hogy a termékek méretkülönbsége 120 bp, tehát valószínűleg jó helyre hibridizáltak a primerek.
2. az RNS 5' végén olyan erősen fel van tekeredve, hogy a reverz transzkriptáz enzim még az mRNS vége előtt abbahagyja az átírást
   - ezt bizonyítja, hogy metil-merkuri-hidroxiddal kezelt RNS-ről hosszabb termékeket tudtunk átíratni.
3. az RNS-nek 5' irányban még hosszú, 500 nukleotidnál hosszabb része van (az irodalomban közölt primer extenziós 5'végpont meghatározásoknál a primer soha sincs messzebb a végponttól 100 bp-nál)
4. magának a kapcsoló fehérje génnek többféle hosszúságú RNS-e van, a transzkripció többszörös kezdőpontról indul.

Mindezeknek a figyelembevételével célszerűnek látszott a keletkezett egyszálú komplementer DNS-t megfelelő vektorba klónozni, hogy így tovább vizsgálhassuk.

3. A klónozás menete

A klónozás során arra törekedtünk, hogy minél hosszabb cDNS-eket kapjunk megfelelő orientációban. Ezért olyan klónozási stratégiát választottunk (7. ábra), amely ötvözi a különböző módszerek előnyeit. Minden cDNS klónozási módszernél kritikus, hogy a képződött cDNS teljes hosszúságban klónozódjon, de mind a hagyományos, az első szál vissza-



8. ábra: Primer extenzió a prE2, és a prE3 primerekről. M: molekulasúly marker; - : metil-merkuri hidroxid kezelés nélkül; + :primer extenzió az RNS metil-merkuri hidroxidos denaturációja után.

hajló végét primerként használó, majd a hurkot S1 nukleázzal felnyitó, mind az RNS-t RN-áz H kezelést követően primerként felhasználó módszereknél elkerülhetetlen, hogy az 5' végről 1-2 vagy több nukleotid leemésztődjön. Egyedül a farkazásos módszerek biztosítják, hogy a megszintetizált első szál teljes hosszában jelen legyen a klónban. Ugyanakkor a homopolimer farokkal ellátott cDNS rosszabb hatásfokkal klónozható, mint a linkerezett. Ezért megpróbáltuk egyesíteni a farkazásos és a többi módszer előnyeit. A cDNS klónozás első lépéseként a prE3 és prE2 oligonukleotidokat független reakciókban hosszabbítottuk

meg. Ezután a cDNS-eket egyesítettük és a továbbiakban együtt kezeltük. Az első szál szintézise után a keletkezett egyszálú komplementer DNS 3' végére egy oligo-(dG) farkat szintetizáltattunk terminális transzferáz enzimmel. A hagyományos farkazásos módszereknél ezután a farkazott cDNS-t farkazott vektorhoz hibridizáltatják, és így folytatják a klónozást. Ennek a hátránya, hogy egyrészt technikailag bonyolult csak egyik végén farkazott vektort előállítani, másrészt a farkazott vektort nagy moláris feleslegben kell alkalmazni, és ez később hátrányosan érinti a második szál szintézisét, és összességében a klónozás hatásfoka alacsony lesz. Mi azonban nem a vektorhoz, hanem egy 8 bp hosszú duplaszálú *Bam*HI linkerhez szintetizáltattunk oligo-(dC) farkat. Ez a farkazott linker szolgált primerként a második szál szintéziséhez, miután a cDNS-en lévő oligo-(dG) farokhoz hozzáhibridizáltuk, majd az RN-áz H-t alkalmazó módszerekhez hasonlóan folytattuk a cDNS szintézist. *Sal*I linker adása után a *Bam*HI és *Sal*I végekkel ellátott cDNS-t azonos enzimekkel hasított pUC12 vektorba építettük be. A rekombináns plazmidokat *E. coli* DH-5 α sejtekbe bejuttatva, összesen mintegy 2000 rekombinánst kaptunk, amelyeket egyenként felszaporítottunk, és glicerin jelenlétében -20 C-on tároltunk.

II. cDNS klónok kiválasztása és jellemzése

1. A kapcsoló fehérje mRNS-sel komplementer klónok kiválasztása

A rekombináns plazmidok közül a pozitívokat kolóniahibridizációval választottuk ki. Próbaként egyrészt a primerként használt két oligonukleotidot, másrészt a pgLP11/8 plazmid (54) *Hin*dIII-*Eco*RI inszertjét, amely a kapcsoló fehérje gén 2. exonját teljes egészében, valamint az azt határoló A és B intronból egy-egy darabot tartalmaz. A pozitív klónokat külön-külön is hibridizáltuk a prE3, valamint az prE2 primerekkel, valamint meghatároztuk a klónozott cDNS-ek méretét. A 2. táblázat foglalja össze annak a 32 cDNS klónnak az adatait, amelyek legalább az egyik oligonukleotid primerrel pozitív hibridizációs jelet adtak. Feltüntettük továbbá az inszert méreteket is. A 32-ből 17 klón csak a prE3 primerrel hibridizál. Ezeknek a klónoknak a többsége nem kapcsoló fehérje cDNS klón, az első szál szintézisekor a primer a mRNS-ek közül valószínűleg, egy nem kapcsoló fehérje gén eredetű hírvívőhöz hibridizált hozzá. Ezt bizonyítandó, kiválasztottuk a pLP8A12 klónt, megjelöltük az inszertet, és csirkmxsternum porc, és koponyatető csont mRNS-t tartalmazó filterhez hibridizáltuk. Eredményként egy olyan méretű nagyon erős jelet kaptunk, mind a sternum, mind a koponyatető mRNS-t tartalmazó filteren, melyeknek a mérete nem egyezik meg egyik kapcsoló fehérje mRNS méretével sem. Összesen 11 olyan klónt találtunk, amely

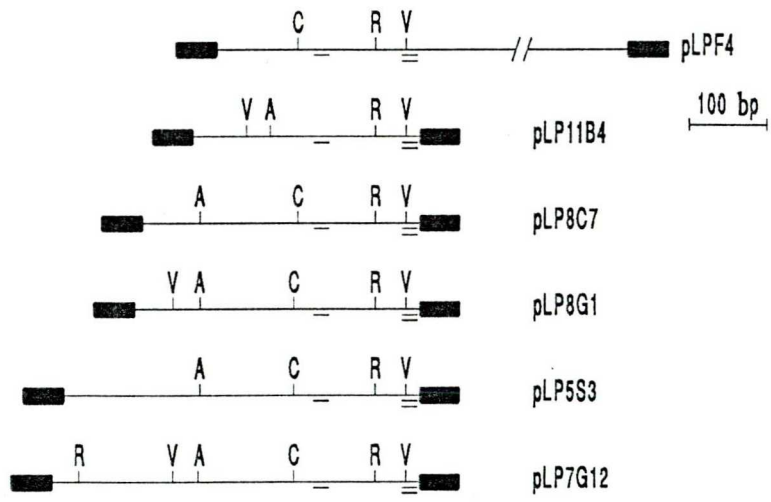
a klón neve	prE3 primer	prE2 primer	11/8 próba	inszert méret
pLP4C10	+	-	-	110
pLP4H11	+	-	-	92
pLP5C2	+	-	-	90
pLP5C5	+	-	-	209,242
pLP5D5	+	-	-	82
pLP7F7	+	-	-	68
pLP8A12	+	-	-	310
pLP9C4	+	-	-	168
pLP9D4	+	-	-	105
pLP9F2	+	-	-	260
pLP10F6	+	-	-	175
pLPE2/16	+	-	-	126
pLPE2/35	+	-	-	86
pLPE1/4	+	-	-	175
pLPE2/1	+	-	-	230
pLP4C9	+	-	-	170
pLPE1/19	+	-	-	129
pLP10F4	+	-	+	175
pLP5F8	+	+	+	239
pLP6D9	+	+	+	340
pLP7G12	+	+	+	520
pLP8C7	+	+	+	405
pLP8G1	+	+	+	430
pLP9F5	+	+	+	283
pLP10G4	+	+	+	238
pLP11B4	+	+	+	339
pLP5S3	+	+	+	498
pLPE2/53	+	+	+	110,139
pLPE1/42	+	+	+	270
pLP10A7	-	+	-	200
pLP7A2	-	+	-	123
pLPE2/59	-	+	-	385

2. táblázat: különböző cDNS klónok adatai. +: az adott próbával pozitív jel, 11/8-as próba: a pgLP11/8 plazmid inszertje (54)

mind a három próbával hibridizál, és mindössze három olyat, amely csak az prE2 primerrel, de később ezek sem bizonyultak kapcsoló fehérje eredetűnek.

2. A cDNS klónok restrikciós térképezése

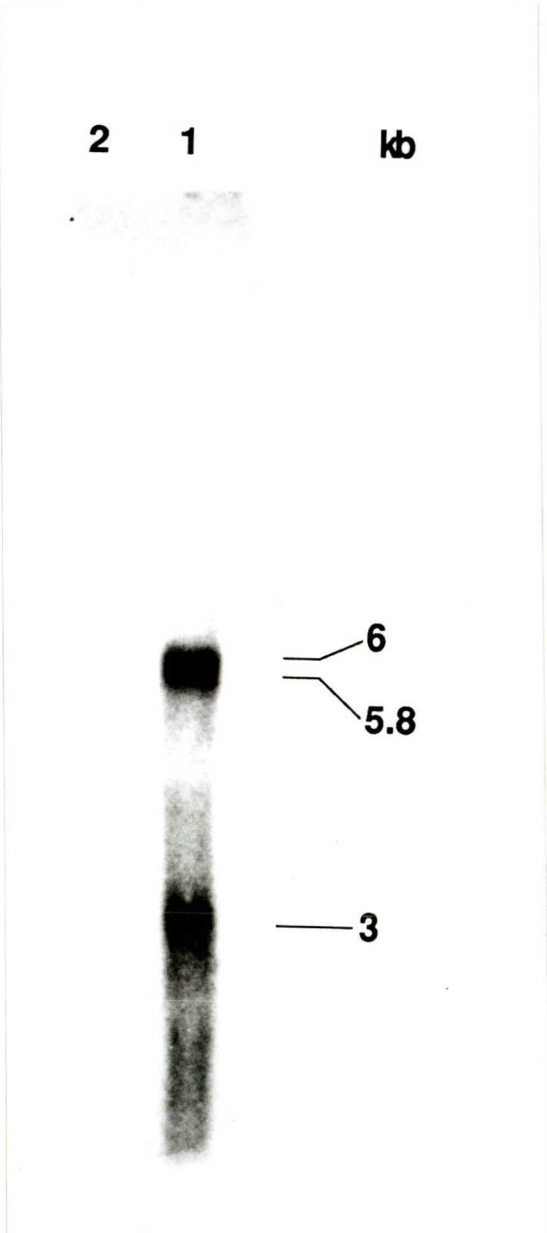
A további kísérletekhez kiválasztottuk a leghosszabb cDNS klónokat, amelyekben meghatároztuk néhány restrikciós endonukleáz hasítási helyét. Így öt cDNS klón restrikciós térképe látható a 9. ábrán. (Mivel a leghosszabb cDNS klón inszertje is csak 480 bp hosszú,



9. ábra: Fontosabb klónok restrikciós térképe a pLPF4 klónhoz (46) viszonyítva. A : *Ava*I, C : *Sac*I, R : *Eco*RII, V : *Ava*II, egyszeres aláhúzás (-): prE2, kétszeres aláhúzás (=): prE3

csak kevés enzim hasítja a klónokat.) Látható, hogy a két primer között teljesen megegyezik mind a hat klón, azonban a prE2 primertől 5′ irányban található *Sac*I hasítóhely hiányzik a pLP11B4 klónból, pedig megtalálható benne az 5′ végen egymástól kb. 30 bp. távolságra két hasítóhely, az *Ava*I és az *Ava*II. Ugyanez a két hasítóhely egymástól hasonló távolságban megtalálható másik két klónban is (pLP8G1, pLP7G12). A pLP8C7, a pLP8G1, a pLP5S3, és a pLP7G12 klónok az *Ava*I helyig megegyeznek egymással a térkép alapján, de ezután a pLP8G1 és a pLP7G12 klónoknál egy *Ava*II hely van, mig a másik két klónban ez a hasítási hely hiányzik. Ezekből az eredményekből azt a következtetést vontuk le, hogy bár az öt cDNS klónban bizonyos heterogenitás tapasztalható, a leghosszabb klón, a pLP7G12, az *Ava*I helyig megegyezik másik 3 klónnal, az *Ava*II helyig pedig a pLP8G1 klónnal, tehát

valószínűleg teljes hosszában a kapcsoló fehérje gén mRNS 5′ végét reprezentálja. Ezt úgy bizonyítottuk, hogy kivágtuk az *Ava*II helytől 5′ irányba terjedő fragmentumot, és ezt használtuk próbaként Northern hibridizációban (10. ábra). A kapott jel egyértelműen megmutatta,



10. ábra: Northern hibridizáció a pLP7G12 klón transzkripciós startpont fele eső részével (az 5′ felén lévő *Ava*II helytől 5′ irányba eső fragmentum). 1: poly (A)+ RNS csirke embriók mellcsont kezdeményéből; 2: poly(A)+ RNS csirke embrió koponyatetőből. Az ábrán be vannak jelölve a kapcsoló fehérje génre jellemző három fajta mRNS méretei.

hogy ez a cDNS szakasz a kapcsoló fehérje gén mRNS-ének átirásából származott, mivel mind a háromféle, korábban meghatározott kapcsoló fehérje hírvivővel hibridizált, és ugyanakkor nem adott jelet a csirke koponyatető mRNS-sel.

3. Az izolált cDNS klónok szekvenciájának meghatározása

A pozitív cDNS klónok további karakterizálása céljából mintegy húsz klónból kivágtuk az inszertet, *Hin*dIII és *Eco*RI enzimekkel végzett kettős hasítással, majd a kapott fragmentumokat M13mp18 és M13mp19 vektorokba klónoztuk, és meghatároztuk a szekvenciájukat.

A szekvenciák vizsgálatából kiderült, hogy a pLP7G12 (és a pLPF4) klónok szekvenciájával átfed, és az átfedő részen teljesen megegyezik a pLP5F8, pLP6D9, pLP9F5, pLP10G4, a pLPE1/42, vagy a pLP10F4 klónok szekvenciája, de ezek közül a pLP7G12 nyúlik legtovább 5' irányban. Csak részlegesen egyezik meg a pLP7G12 klón szekvenciájával a pLP11B4, a pLP8G1, pLP8C7 és a pLP5S3, amint az már a térképezésből is kiderült (9. ábra). Ezekkel a klónokkal a későbbiekben még foglalkozunk.

Bebizonyosodott, hogy mind a prE2 primerről indult 3 klón, mind a csak a prE3 primerrel hibridizáló klónok valóban hamis pozitívok, ezekben a primer szekvenciáján kívül más kapcsoló fehérje DNS-sel homológ szakaszt nem találtunk.

Összeségében 32, a kolónia hibridizáció során valamelyik primerrel pozitív klón közül 14 bizonyult kapcsoló fehérje eredetűnek. Tehát az összes (2000) cDNS klónnak több mint fél százaléka.

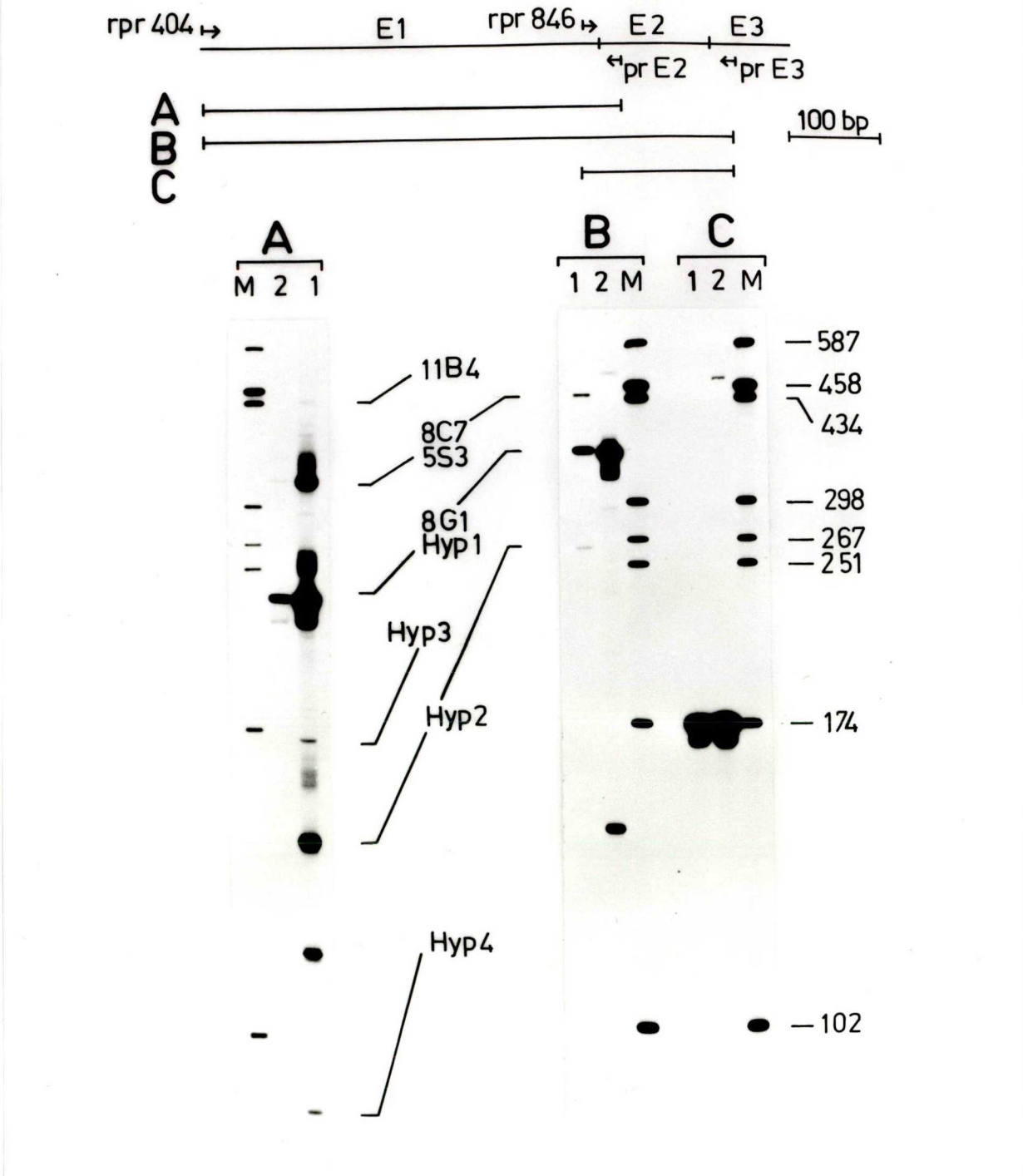
III. Alternatív exon keresés a második és harmadik exonok között

1. A pozitív cDNS klónok közötti keresés

A prE3 primert azért is használtuk, hogy klónozni tudjuk, ha létezik, a második és harmadik exon között elhelyezkedő exont, amely homológ a patkány kondroszarkoma kapcsoló fehérje génnél kimutatott alternatív exonnal (55). Az átvizsgált 29 prE3 primerről indult cDNS klón között azonban nem találtunk olyat, amelyikben benne lenne az alternatív exon, igaz, ezeknek a klónoknak több, mint a fele nem kapcsoló fehérje eredetű, de a fennmaradó 14 klón között is valószínűsíthető volt, hogy egy alternatív exont hordozó, ha elfogadjuk a Rhodes és mti. (55) által megállapított 1:10-hez arányt az alternatív exont hordozó és nem hordozó mRNS-ek között.

2. Polimeráz láncreakció az alternatív exon kimutatására

A klónozás eredményét nem tekintettük végleges bizonyítéknak, ezért egy érzékenyebb kísérletet is elvégeztünk a csirke kapcsoló fehérje génben a 2. és 3. exonok között esetleg előforduló alternatív exon létének kimutatására. Szintetizáltattunk egy reverz



11. ábra: PCR termékei denaturáló poliakrilamid gélen elválasztva, autoradiografálva. A prE2 és a prE3 primerek 5′ végét jelöltük rádioaktiv foszforral. M: molekulasúly markerek; 1. oszlopok: PCR poly (A)+ RNS-ről átirt komplementer DNS templáton; 2. oszlopok: kontroll PCR pLP8G1 plazmid DNS templáton. Középen a különböző cDNS formák láthatóak. Magyarázat a szövegben.
primerpárok; A: rpr404 és prE2
B: rpr404 és prE3
C: rpr846 és prE3

primert, amely az első exon 3' végével homológ, és ezzel, valamint a prE3 primerrel PCR-t végeztünk. Amennyiben akár az első és a második, akár a második és a harmadik exonok között alternatív exont hordozó mRNS molekula létezik, akkor a PCR során annak is amplifikálódni kell. A kísérlet eredményeképp azonban nem tudtunk kimutatni a normális 174 bp-nál hosszabb terméket, pedig a detektálás érzékenységét rádioaktív izotóp használatával növeltük (11. ábra C). Tehát a csirke kapcsoló fehérje génben nem találtunk alternatív exont az első és a harmadik exonok között.

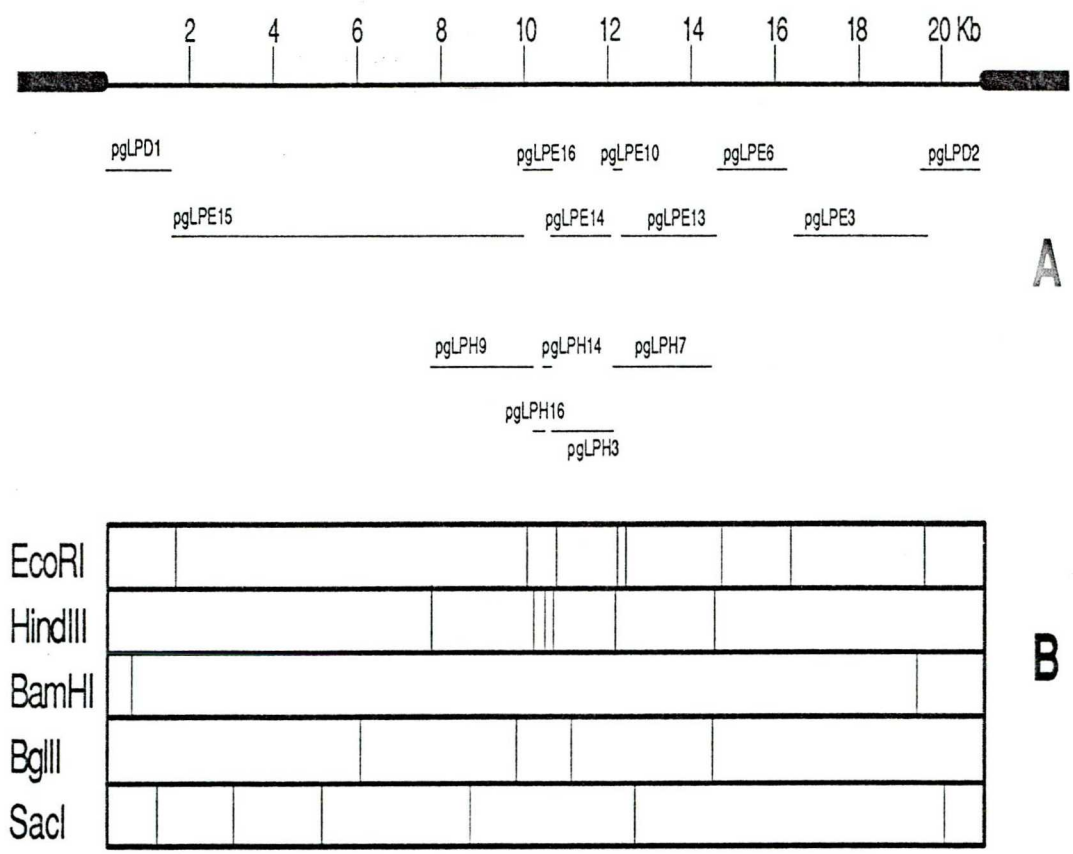
IV. Az első exont tartalmazó genomiális klón

Az izolált cDNS klónok közül a pLP7G12 nyúlik legtovább 5' irányban, és mint az előzőekben bebizonyosodott (10. ábra), végig kapcsoló fehérje eredetű. Ezért ez a klón látszott a legalkalmasabbnak arra, hogy próbaképp használjuk az első exon és a promoter régió génkönyvtárból való kikereséséhez. Azért, hogy csak specifikusan az 1. exonnak megfelelő jelet kapjunk, átklónoztuk a pLP7G12 klón *Sac*I helytől 5' irányba eső részét (a *Sac*I hely csak néhány bázispárnyira van az első és a második exon határától), és ezt a fragmentumot jelöltük meg a plakkhibridizációhoz. Silvija Mestric a jelölt fragmentum segítségével izolált egy pozítiv klónt a csirke genomiális klónkönyvtárból (54). A klón neve λgLP100 és benne a csirke DNS 21 kb hosszú.

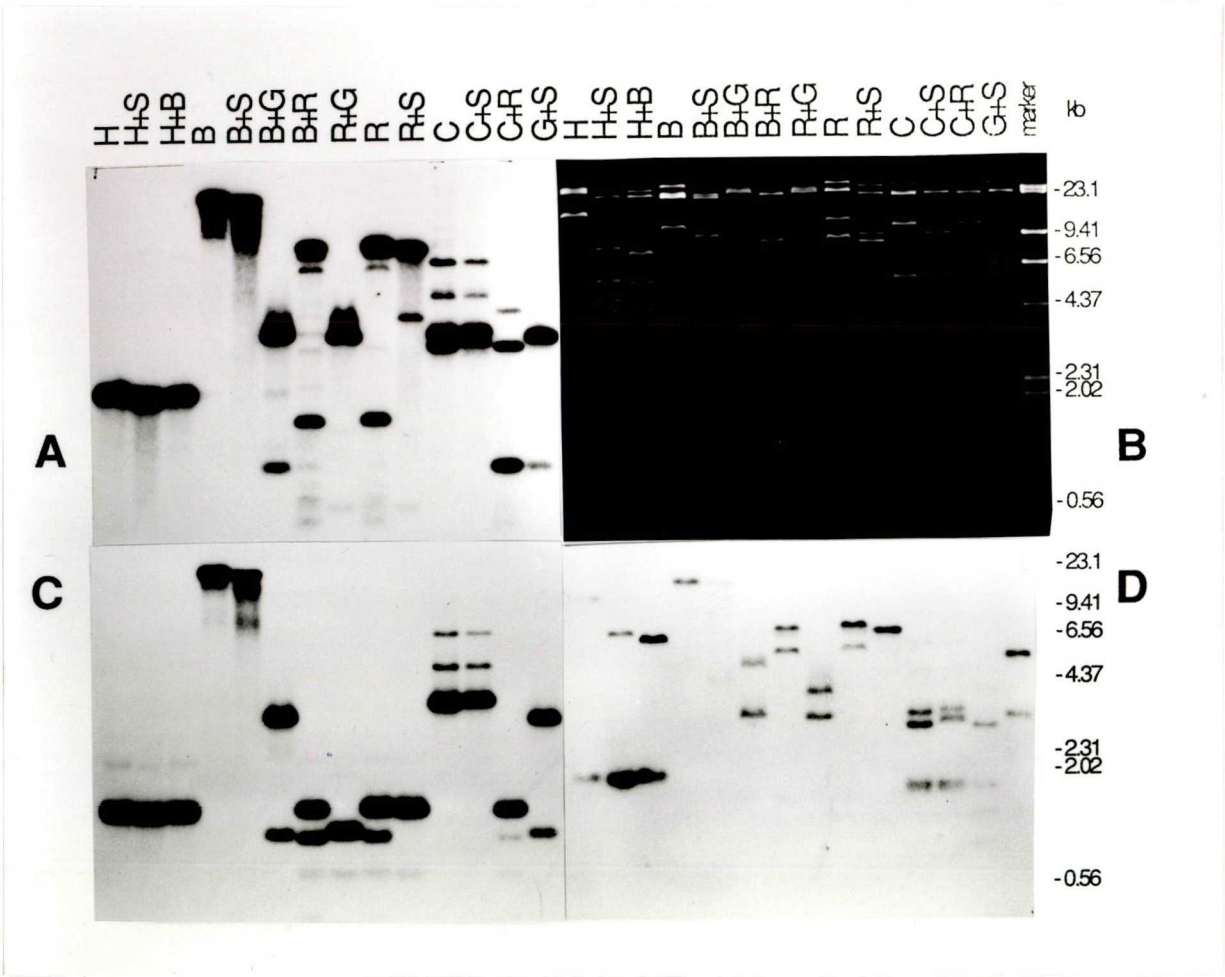
1. A λgLP100 klón térképezése

A λgLP100-as klónt felszaporítottuk és szubklónoztuk belőle a *Hin*dIII, *Eco*RI, és az inszert két szélén található *Sal*I- *Eco*RI fragmentumokat (12. ábra A). A kapott szubklónoknak meghatároztuk a méretét és a restrikciós térképét. Egyes klónokat megjelöltünk, és próbaként használtunk Southern hibridizációban (13. ábra), hogy megállapítsuk a λgLP100 teljes fizikai térképét 5 restrikciós endonukleázra (12. ábra B). Meg kell említeni ezenkívül, hogy számos enzim (*Sal*I, *Xho*I, *Xba*I, *Sma*I,) a klón viszonylag nagy mérete (21 kb) ellenére, egyetlen helyen sem hasítja azt.

Az első exont hordozó próba a λgLP100 jobb kar felé eső végéhez hibridizál, az 1.7 kb-os pgLPD1 klón tartalmazza a teljes első exont. Ez egyben azt is jelenti, hogy a klónozott régió nagy részét (19.7 kb), az első exon után következő első intron teszi ki, és a feltételezett promoter régióból is tartalmaz mintegy 900 bp-t. A λgLP12.1 klón (54) ugyancsak tartalmaz az első intron 3' végéből egy 13.6 kb-os darabot. A két klón azonban restrikciós térképeik alapján nem fed át, és nem is hibridizálnak egymással. Ebből következik, hogy az első intron hosszabb mint 33 kb, a kapcsoló fehérje génje pedig több mint 100 kb-t foglal el a genomban!



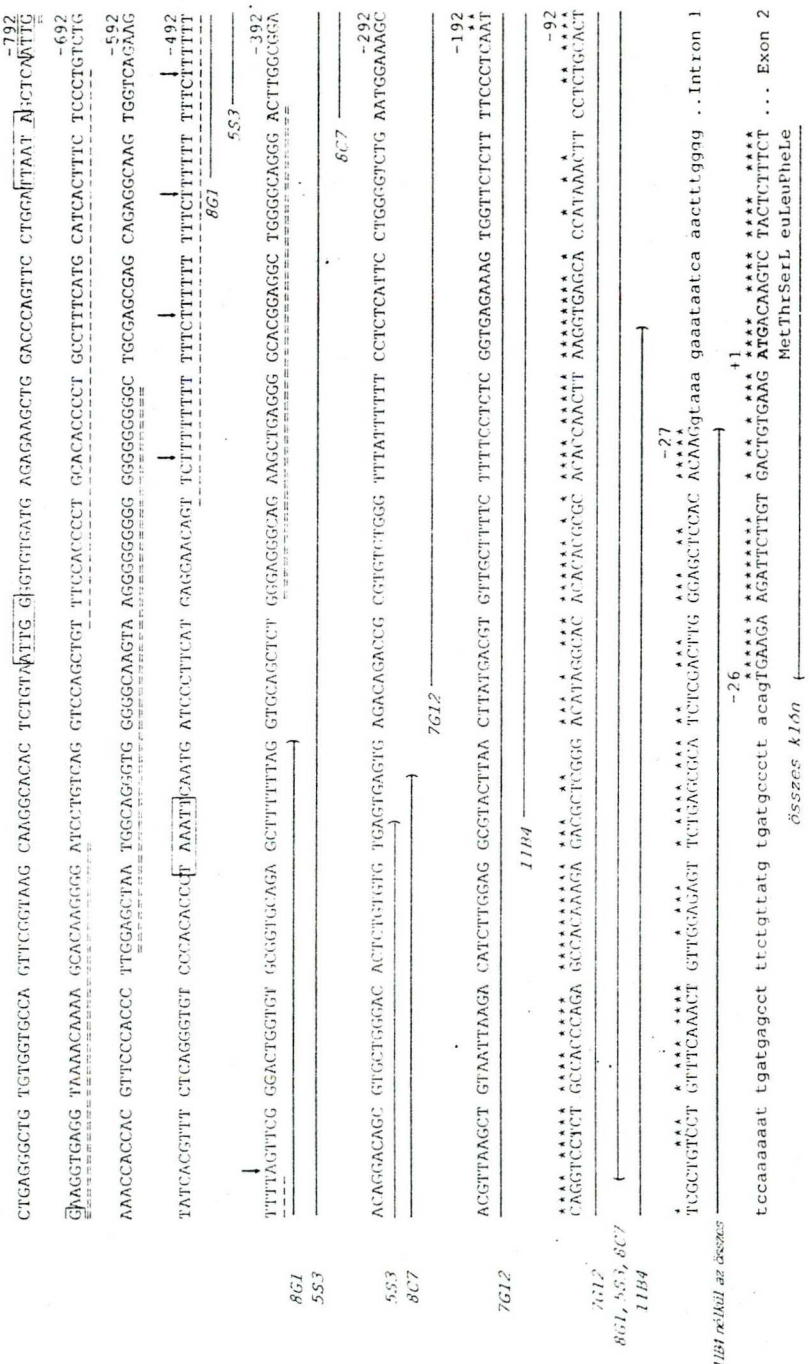
12. ábra: a λgLP100 klón szubklónjai, és restrikciós térképe. A: a pgLPD1 és pgLPD2 az inszert két szélén található *Sal*I-*Eco*RI fragmentumokat, a pgLPEx az *Eco*RI fragmentumokat, a pgLPHy a *Hind*III fragmentumokat tartalmazzák. B: a λgLP100 klón inszertjének térképe ötféle restrikciós enzimmel.



13. ábra: a λgLP100 klón Southern hibridizációs vizsgálata. B: λgLP100 fág DNS különböző restrikciós endonukleázokkal hasítva (H: *Hin*dIII; S: *Sal*I; B: *Bam*HI; G: *Bgl*II; R: *Eco*RI; C: *Sac*I);
próbák: A: pgLPH9,
C: pgLPH3, és
D: pgLPE15 (lásd 12. ábra)

2. A pLPD1 klón

Mint már említettük, a pLPD1 szubklón tartalmazza az első exon próbával hibridizáló szakaszt. Ennek a klónnak részletesebben is meghatároztuk a restrikciós térképét. A térkép alapján további szubklónokat készítettünk, majd meghatároztuk az első exon és az azt határoló szakaszok DNS szekvenciáját (14. ábra).



14. ábra: A csirke kapcsoló fehérje gén 5' végének a szekvenciája. +1: a transzláció kezdőpontja a második exonban. A cDNS klónok folyamatos vonallal vannak aláhúzva. A 30 nukleotidnál hosszabb purin és pirimidin gazdag szekvenciák kettős, illetve sima szaggatott vonallal vannak jelölve. A TATA és CCAAT box szekvenciák be vannak keretezve. A transzkripció fő kezdőpontjait nyilak jelzik. csillag jelöli a megegyezést az ember kapcsoló fehérje cDNS szekvenciájával (53).

V. Alternatív splicing a kapcsoló fehérje mRNS 5' végi nem transzlálódó részén

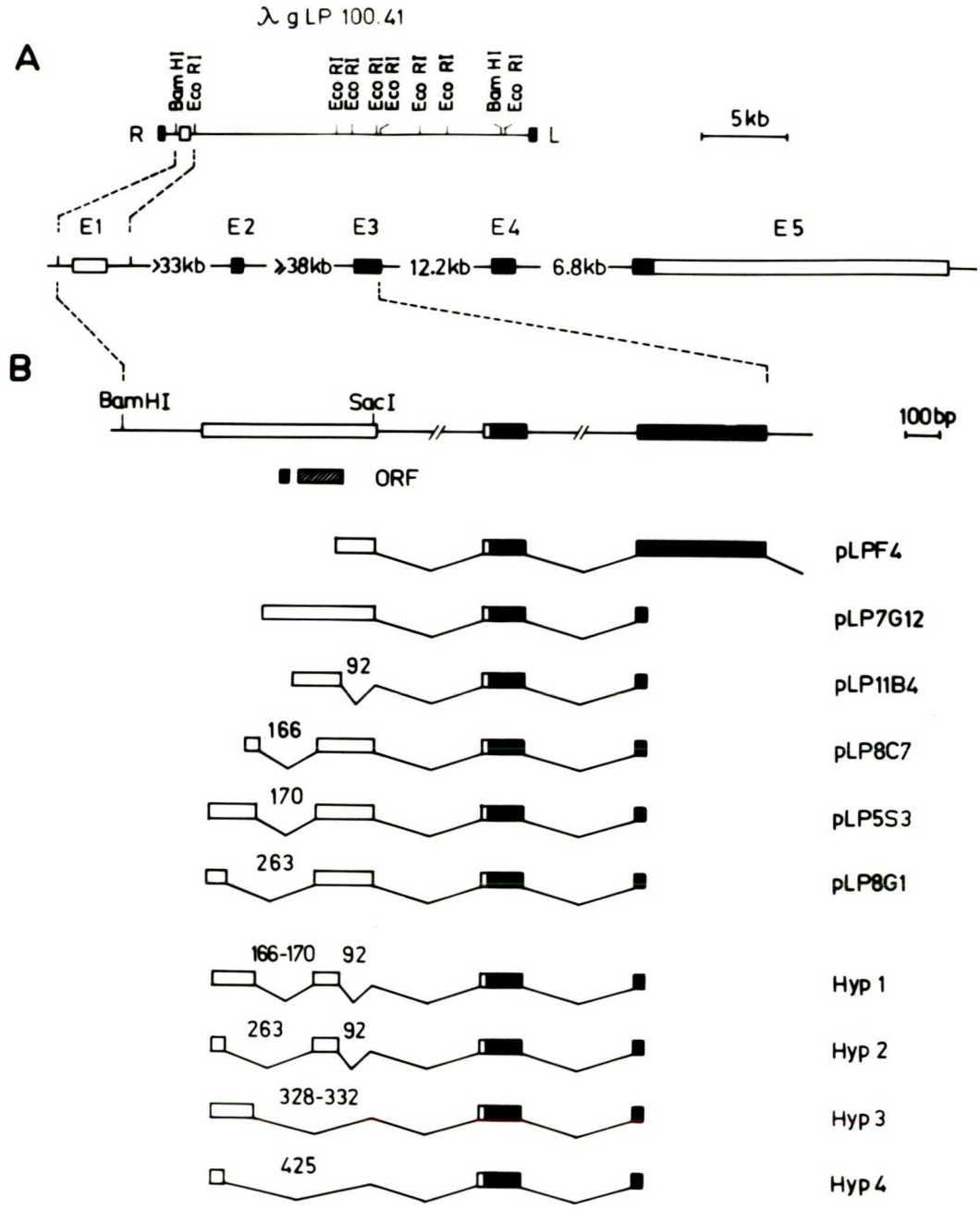
Mivel a genomiális klón keresésekor a pLP7G12 klón első exont tartalmazó szubklónját használtuk, először ennek szekvenciáját hasonlítottuk össze a genomiális pgLPD1 nukleotid sorrendjével. Nem okozott nagy meglepetést, hogy azok az átfedő részen teljesen megegyeztek egymással. Az összehasonlításból meg lehetett állapítani az első intron kivágódási donor helyét (3. táblázat) is. Mivel a többi cDNS klónnak is meghatároztuk már a nukleotidsorrendjét és tudtuk, hogy ezek restrikciós térképeiket (9. ábra) és szekvenciáju-

3. táblázat, Feltételezett kivágódási helyek a kapcsoló fehérje gén 5' végén

donor hely[a]	poz[b]	akceptor hely[c]	poz[b]	cDNS klón
		ttttttttagTTC	-485	
		gctttttagGTG	-451	
TAGgtgcag	-451	cctcaatcagGTC	-188	pLP8G1
TGTgtgagt	-362			
TGAgtgagt	-358	cctcaatcagGTC	-188	pLP5S3
TGAgtgaga	-354	cctcaatcagGTC	-188	pLP8C7
TCGgtgaga	-221			
AAGgtgagc	-118	gcccttacagTGA	-26	pLP11B4
AAGgtaaag	-27+1	gcccttacagTGA	-26	pLP11B4 kivételével az összes
C/A AGgtragt		yyyyyynyagG		konszenzus[d]

[a] Motívumok, melyekben megtalálható az AGgtr vagy gtrag szekvencia. r= a vagy g.
[b] A kivágódási hely pozíciója a 14. ábra szerint.
[c] Motívumok, melyek az aláhúzott részen megegyeznek a konszenzussal, és a többi helyen sem különböznek attól többel, mint két nukleotid. y= c vagy t.
[d] Shapiro és mti. (98)

kat tekintve is eltérnek egymástól és a pLP7G12 klóntól, ezek szekvenciáját is összehasonlítottuk a genomiális klónéval. A várakozással ellentétben azonban ezek szekvenciája is teljes hosszúságban megegyezik a genomiális klónéval, azzal a különbséggel, hogy a cDNS-ekből hosszabb-rövidebb szakaszok hiányoznak! A "normális" pLP7G12 klónon kivül 4 féle változatot találtunk; a pLP11B4 92, a pLP8C7 166, a pLP5S3 170, és a pLP8G1 263 nukleotid hosszúságú deléciót tartalmaz (15. ábra). Mi lehet ennek a jelenségnek magyarázata?



15. ábra: Különböző cDNS formák a kapcsoló fehérje génhez viszonyítva. A: a λgLP100 klón az első exonnal (üres keret); B: A keletkezett 5 féle cDNS forma, a számok a genomiális DNS-hez képest kivágódott szekvencia hosszát jelentik. HYP1 - HYP4: a PCR során keletkezett hipotetikus termékek (11. ábra A és B), ORF: a bevonalazott téglalapok az első exonban levő nyitott leolvasási keretek pozicióját és hosszát jelentik.

Ismert, hogy egyes géneknél az elsődleges transzkriptumból, a pre-mRNS-ből többféle érett hírvivő is keletkezhet. Ez a jelenség az "alternatív splicing", mikor a pre-mRNS érése során az intronok kivágódása alternatív helyeken történik. Feltételezzük, hogy ez történik a kapcsoló fehérje gén 5' végén is, a deléciók tulajdonképpen egy-egy intronnak felelnek meg, kivágódásuk *in vivo* is lejátszódik, és mivel egy DNS szakaszról többféle mRNS is keletkezik, alternatív splicing történik.

Feltételezésünk igazolására a következő bizonyítékokat szereztük:

1. Konszenzus kivágódási helyek megléte

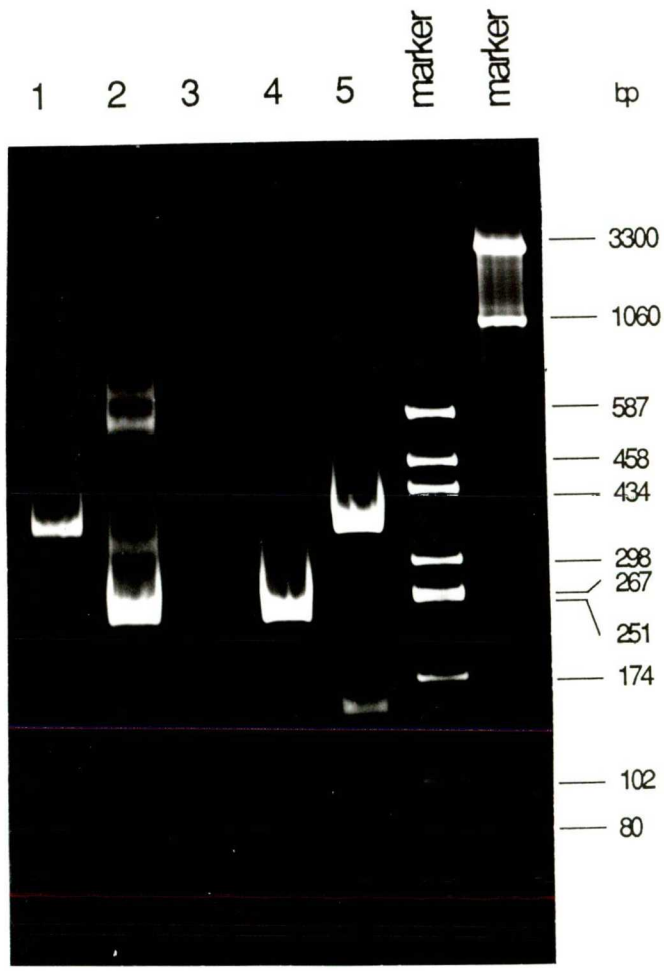
Megvizsgálva a kivágódott intronok és a velük szomszédos exonok végeit kiderült, hogy mindegyik megegyezik, vagy csak egy-két nukleotid eltérést mutat a kivágódási donor és akceptor helyek konszenzus szekvenciájával (98) (3. táblázat). Ez azt jelenti, hogy a splicing ma ismert legfontosabb kritériumának, a megfelelő donor és akceptor szekvenciák meglétének megfelelnek a deléciók. Ráadásul ha azt nézzük, hogy az adott régióban hány lehetséges kivágódási hely van, akkor azt tapasztaljuk, hogy a 7 lehetséges donor helyből 5, és a 4 lehetséges akceptor helyből 2 hasznosul ténylegesen is. Tehát arra következtettünk, hogy a cDNS klónokban talált deléciók tulajdonképpen a pre-mRNS érése során kivágódott alternatív intronok, melyeket normális kivágódási akceptor és donor helyek határolnak.

2. Alternatív splicing bizonyítása polimeráz láncreakcióval (PCR)

Az 5' nemtranszlálódó részen történő alternatív splicing *in vivo* meglétére további bizonyítékot szolgáltattak a cDNS szintézishez kapcsolt polimeráz láncreakció (PCR) kísérletek eredményei. Csirke mellcsont kezdeményből izolált poly(A)$^+$ RNS-ről a prE3 és prE2 primerek meghosszabításával egyszálú cDNS-t írattunk át, ugyanúgy, mint a primer extenziónál. Az RNS elhidrolízálása után, hozzáadtuk a PCR primereket, 5' irányból az rpr404-et, 3' irányból pedig vagy a prE3 vagy a prE2 primert. Kontrollként a pLP8G1 plazmid linearizált formájából 1 ng DNS-t használtunk ugyanezekkel a primerekkel (11. ábra A és B). Az autoradiogramon a 2. oszlopokban láthatóak a kontroll DNS-en, az 1. oszlopokban pedig az RNS-ről átírt cDNS-en végzett PCR eredményei. Legerősebb jelet a 8G1 típusú termék adja, kissé halványabb a 8C7 és az 5S3 típusú termékek csíkja. A prE2 primerrel végzett reakcióban halványan sejteni lehet a 11B4 típusú terméket is. Nem lehet azonban látni a 7G12 klónnak megfelelő terméket. Ezt lehet magyarázni azzal a ténnyel, hogy a reakció során előnyben vannak a rövidebb termékek, de az is lehetséges, hogy valóban nagyon kevés van ebből a cDNS formából. Ezen típusok mellett fellelhetők még az

ún. hipotetikus formáknak megfelelő csíkok is. Ezeket a formákat a klónozott cDNS klónoknál található kivágódási helyek kombinálásával lehet levezetni (15. ábra). A hipotetikus formák között legerősebb a HYP2 formának megfelelő csík. Ugyanakkor ezek a formák annak ellenére, hogy rövidebbek gyengébb jelet adnak, mint a 8G1 típusú mRNS tehát abból sokkal többnek kell lennie.

Hogy tovább erősítsük a gyenge csíkokat, kivettük a prE3 és az rpr 404 primerekkel végzett PCR egy részét és újabb 30 ciklust végeztünk vele (16. ábra). A reakció specifitását



16. ábra: PCR eredménye poliakrilamid gélen futtatva etidium-bromiddal festve. 1, 2, 3 : csirke mellcsont kezdeményből izolált poly (A)+ RNS-ről átírt egyszálú cDNS templáton kétszer 35 ciklus; 4,5 : pLP8G1 linearizált plazmid DNS templáton 35 ciklus.
primerpárok: 1, 5: rpr404 és prE2
2, 4: rpr404 és prE3
3: negatív kontroll.

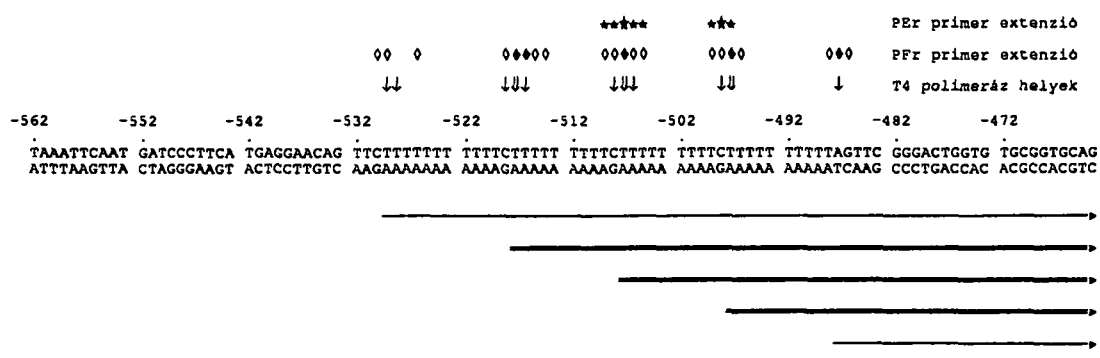
tovább növelte amikor az eredetileg a prE3 primer meghosszabbításával nyert cDNS-en prE3 primerrel végzett láncreakció termékét prE2 primerrel sokszorosítottuk tovább A második amplifikáció eredményeképpen már elég termék keletkezett ahhoz, hogy etidium-bromid fluoreszcencia alapján is látható legyen (16. ábra). Itt is a 8G1 típusnak megfelelő csík a legerősebb, de látható a 8C7 és az 5S3, a 11B4 tipusú forma is. A hipotetikus alakok közül a HYP2 forma látszik. A magasabb molelulasúly tartományban látható csíkok valószínűleg különböző PCR melléktermékeket, heteroduplex molekulákat reprezentálnak.

VI. A transzkripció kezdőpontjának meghatározása

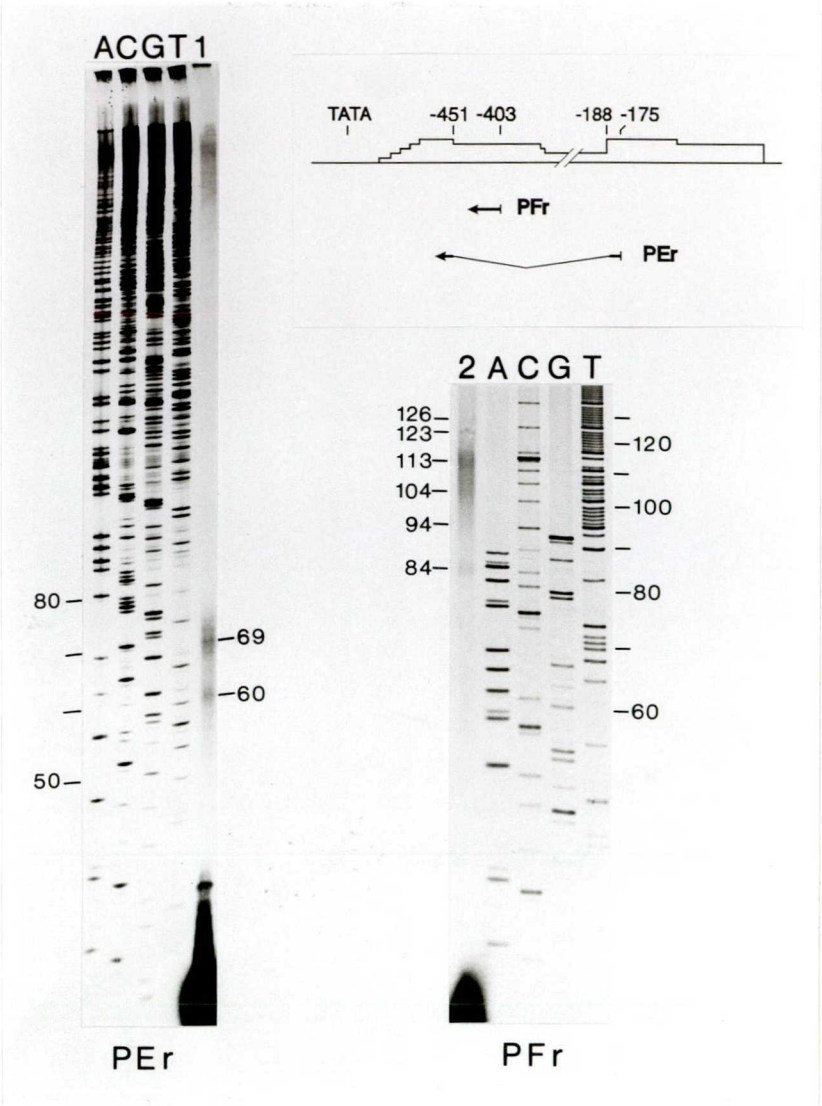
Korábban már említettük, hogy az eredetileg használt prE3 és prE2 primerek hibridizálási helye túl messze van a mRNS 5' végéhez képest, így a primer extenzióval nem tudtuk megállapítani pontosan, hogy mekkora szakasz van még 5' irányban, csak azt tudtuk mondani, hogy a prE2 primerhez képest is minimum 400-500 bp. Az első exon klónozásával és szekvenciájának meghatározásával lehetővé vált, hogy az 5' véghez közelebbi primereket (PEr, PFr, 1. táblázat) használjunk a primer extenzióhoz. Mivel az exont és a feltételezett promoter régiót tartalmazó genomiális DNS szakasz is a rendelkezésünkre állt, a transzkripció iniciációját T4 polimerázos térképezéssel is meg tudtuk határozni.

1. Primer extenzió a PEr és a PFr primerekkel

A primer extenziót az előzőekben leírt módon végeztük. A szintetizált egyszálú, jelölt DNS-t szekvencia létra marker mellett futtattuk, hogy megállapíthassuk a termékek pontos méreteit. A 17. ábrán látható, hogy az általánossal ellentétben a mRNS szintézis nem egyetlen, hanem több, egymástól 9-10 bp-nyira lévő pontról indul, méghozzá a T gazdag régióból. A PEr primer meghosszabításával két fő termék keletkezik, 60 és 69 bp hosszúság-



17. ábra: A transzkripció kezdőpontjai a szekvenciához viszonyítva. A vastagabb vonalak a gyakrabban használt iniciációs pontokból indulnak ki.



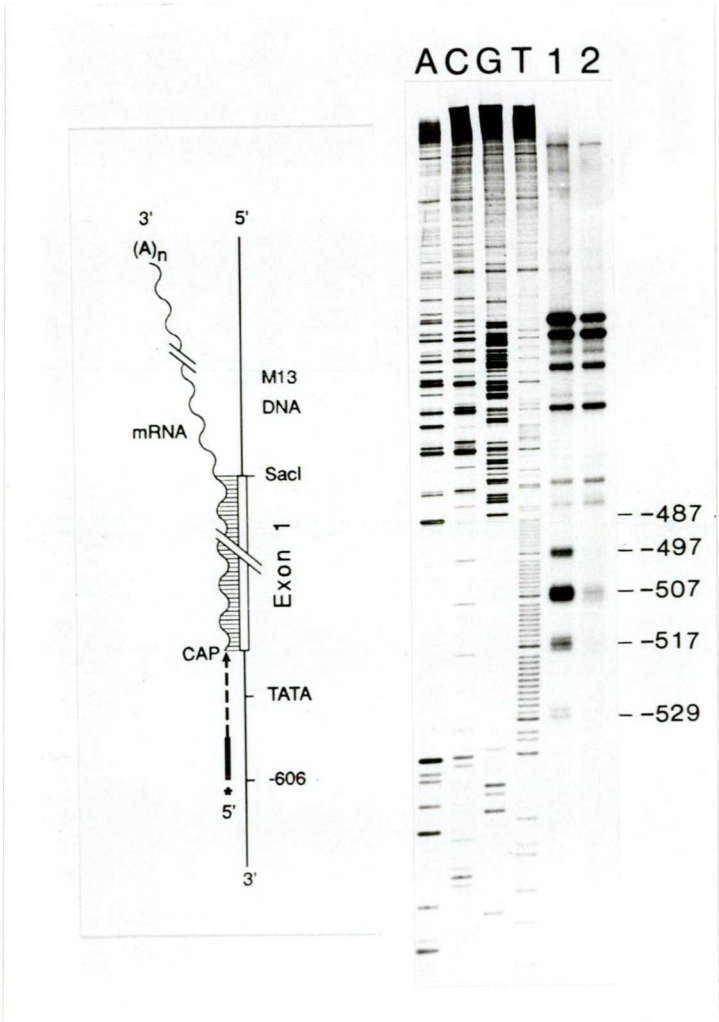
18. ábra: primer extenzió a PEr és a PFr primerekkel, denaturáló poliakrilamid gélen elválasztva, szekvencialétra mellett.

ban (18. ábra). Mivel ezt a primert úgy terveztük, hogy specifikusan a 8G1-es típusú transzkriptumhoz hibridizáljon (a közepére esik az a hely, ahonnan kivágódik egy intron), a kapott termékeket úgy értelmezhetjük, mint a 8G1 tipusú mRNS-ek 5' végi darabjainak a komplementerét.

A PFr primer a 8G1 tipusú transzkriptum kivételével az összes fajta mRNS-hez hibridizálhat. Meghosszabbításával 5-6, egymástól 9-10 bp-ral eltérő fragmentumot kaptunk. Ezeknek a vége szintén az előbb már említett T gazdag régióra térképeződik, és magában foglalja a PEr primerrel kapott iniciációs helyeket is (18. ábra).

2. Transzkripciós startpont meghatározása T4 polimerázzal

A transzkripciós kezdőpontok pozíciójának pontosabb meghatározása céljából egy másik kísérletet is elvégeztünk. A pgLPD1 klón egy M13 szubklónjából, amely tartalmazza a feltételezett promoter régiót és az 5′ nem transzlált részt, az mRNS-sel komplementer egyszálú DNS-t izoláltunk. Ehhez hozzáhibridizáltuk az mRNS-t, és egy 5′ végén jelölt primert, amely a feltételezett promoter régióhoz hibridizált. A primert T4 polimerázzal meghosszabbítottuk. Azt vártuk, hogy a polimeráz csak addig tud haladni az egyszálú DNS templáton, amig bele nem ütközik az odahibridizált mRNS 5′ végébe. Ugyanezzel a primerrel, ugyenezen az egyszálú templáton egy szekvenciázó reakciót is végeztünk, és együtt futtattuk a T4 polimerázzal átírt DNS-sel. A 19. ábrán nagyon jól látható, hogy a polimeráz



19. ábra: A transzkripciós startpont T4 polimerázos térképezése. 1. oszlop: primer extenzió előtt RNS hozzáhibridizálásával; 2. oszlop: RNS hozzáadása nélkül

terminációja, ami megfelel az mRNS iniciációjának, szintén a T gazdag régióban történik, és összhangban van a primer extenzióval kapott pontokkal (17. ábra).

VII. A genomiális DNS szekvencia számítógépes elemzése

Mind a primer extenzió, mind a T4 polimerázos térképezés, valamint az a tény is, hogy egy cDNS klón sem nyúlik tovább, azt tanúsítja, hogy a kapcsoló fehérje gén mRNS transzkripciója a 14. ábra szerinti -361-től -407-es poziciótól terjedő T gazdag régióban kezdődik. Az ettől 5' irányban terjedő genomiális DNS szakasz tekinthető tehát a gén promoterének. Mivel ez a szakasz a λgLP100 klónnak a jobb kar felőli végén található, összesen mintegy 860 bp áll rendelkezésünkre a promoter régióból.

A megszekvenált régióra, amely magában foglal 360 bp-t a promoter régióból, az első exont, valamint néhány bp-t az első intronból, jellemző, hogy benne hosszabb (30-50 bp) purin (kettős szaggatott vonallal aláhúzott), valamint pirimidin (szaggatott vonallal aláhúzott) gazdag régiók váltják egymást (14. ábra).

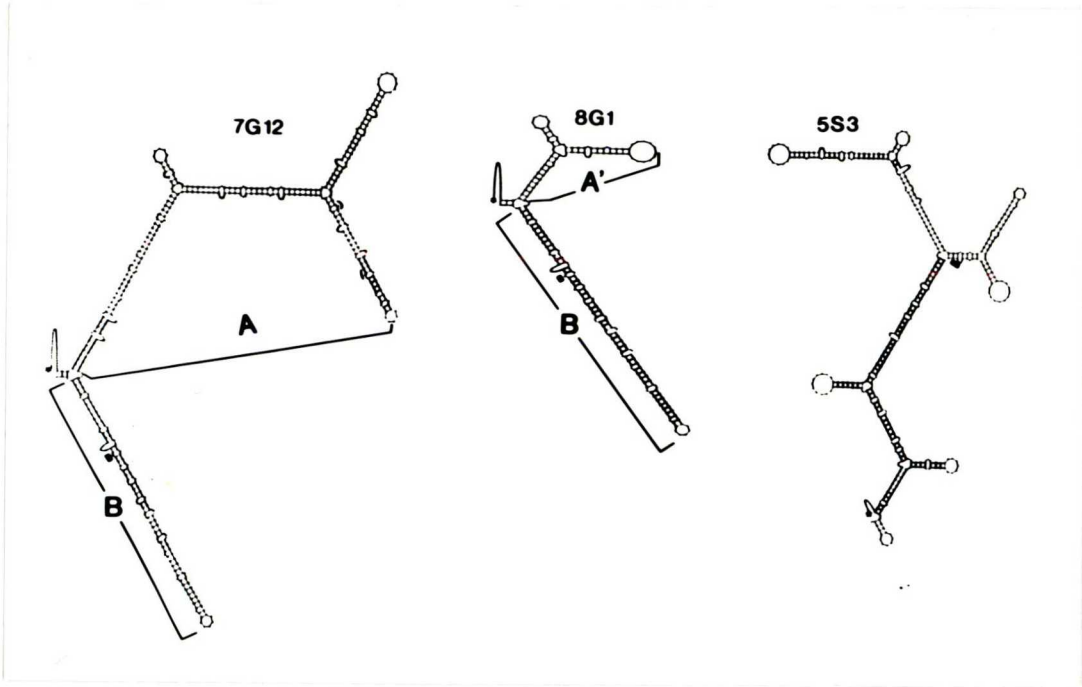
Két TATA boxra hasonlító motívumot találtunk (a 14. ábrán bekeretezve). A primer extenzióval és a T4 polimerázos térképezéssel megállapított transzkripciós kezdőpontok a közelebbi TATA boxtól 34±1, 46±1, 56±1, 66±1 és 76±1 bp-nyi távolságra helyezkednek el, és jellemző rájuk, hogy mindössze egy pontban (aláhúzva) térnek el az iniciációs pontokra megállapított YYCAYYYY (Y= C vagy T) konszenzus szekvenciától (99). Két CCAAT boxot is találtunk, ezek 265 és 315 bp távolságra vannak az első kezdőponttól.

Megvizsgáltuk számítógépes összehasonlító programokkal, hogy található-e közös szekvencia motívum a kapcsoló fehérje gén, és az ugyancsak a porcban található porc mátrix fehérje gén promoter régiója között (28). Az kapott eredmények azt mutatják, hogy nem található ilyen régió az általunk vizsgált szakaszon.

Az 5' nem transzlált régió teljes hosszúsága a csirke kapcsoló fehérje génje esetében több, mint 500 bp. Marilyn Kozak összehasonlító vizsgálatai során (100) 346 gerinces gén közül mindössze négyet talált, amelyek 500 bp-nál hosszabb 5' végi nem transzlált régióval rendelkeztek.

Érdekesség, hogy az 5' nem transzlált régióban két extra AUG, azaz transzláció iniciációs kodon található. Az első egy nyolc aminosavból álló, míg a második egy 42 aminosavból álló nyitott leolvasási keretet határoz meg. Az alternatív splicing miatt azonban ez a két upstream AUG kódon csak a 11B4 és a 7G12 formájú transzkriptumokban szerepel, a többi típusúból kivágódik.

Megvizsgáltuk, hogy a különböző típusú transzkriptumok 5' végei mennyire képesek



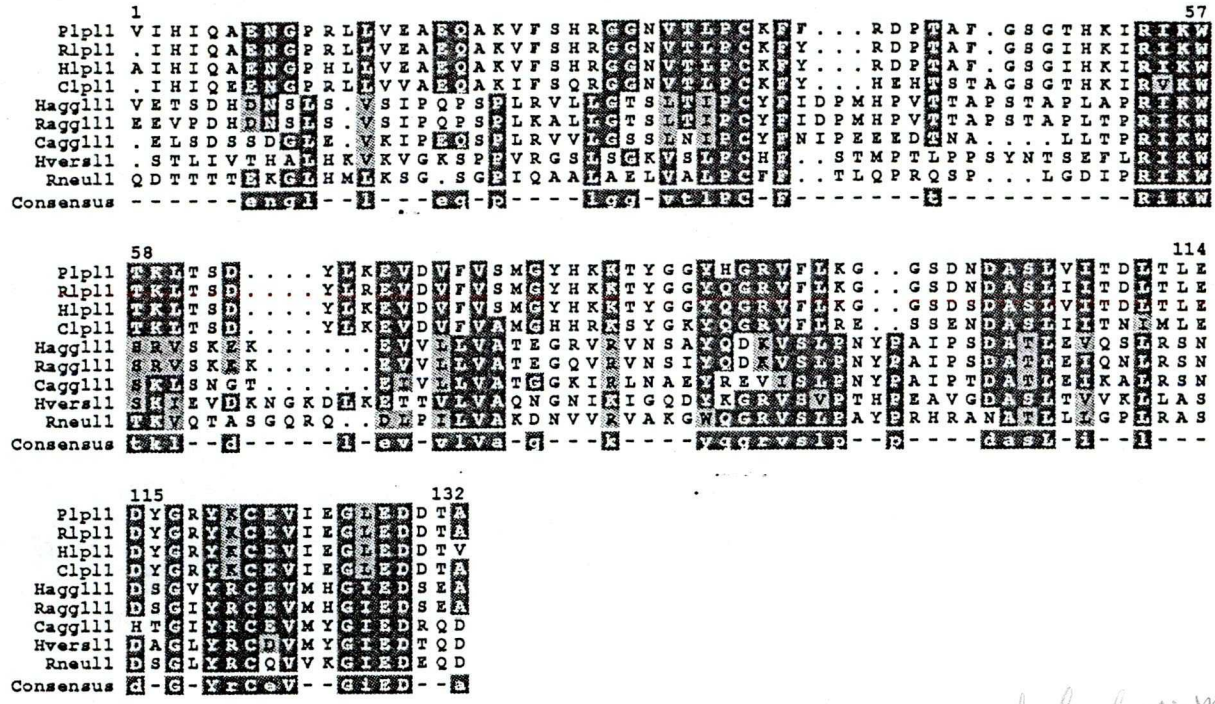
20. ábra: Három cDNS 5' végének másodlagos szerkezet predikciója Zuker algoritmusa (96) szerint a GCG programcsomagban található Fold programmal. Vastag vonal a transzláció kezdőpontját, vékony vonal az 5' nem transzlált régióban levő AUG kodonokat jelzi

önmagukkal párosodva hajtűhurkokat alkotni. A komputer által számított másodlagos szerkezet predikciók láthatók a 20. ábrán. A leghosszabb transzkriptum (7G12) esetében két nagy hajtű keletkezik (A, B), míg a 8G1 típusúnál ebből csak a B marad meg , az A helyett egy rövidebb (A') keletkezik. Ugyanakkor az 5S3 típusú RNS-nél az A és a B fajtájú hajtű sem található meg.

VIII. A kapcsoló fehérje gén molekuláris evolúciója

1. Az első hurok

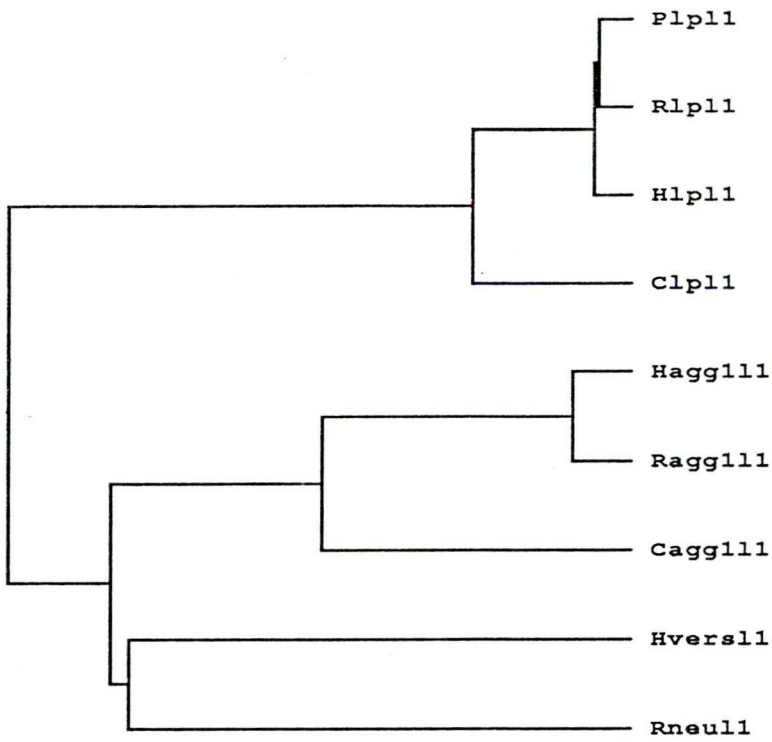
A csirke (46), ember (52), sertés (52), és patkány (50) kapcsoló fehérjékben, az emberi (64), csirke (65) és a patkány (18) aggrecanban, az emberi versicanban (66) és a patkány neurocanban (67) található első hurok aminosav szekvenciákat különböző multiple alignment programokkal hasonlítottuk össze. A hurok határait a rendelkezésünkre álló csirke és patkány kapcsoló fehérje gén exonhatárok alapján állapítottuk meg. A programok által generált összehasonlítások eredményeit kézzel tovább finomítottuk. A 21. ábrán láthatóak az első hurok szekvenciák egymáshoz illesztve, és a konszenzus szekvencia. Az összehasonlított szekvenciák egymáshoz viszonyított eltérései alapján a PileUp program se-



21. ábra: a rendelkezésre álló első hurok szekvenciák összehasonlítása és a levonható konszenzus szekvencia. A szekvenciák elnevezése a következőképpen történt: A végén az l1 arra utal, hogy első hurok szekvenciáról van szó, az lp kapcsoló fehérjét, az ag az aggrecant, a vers a versicant, a neu a neurocant jelöli. A g1 az aggrecan G1 globuláris egységét jelenti. A kezdő betű adja meg a faj nevét, a P a sertés, R a patkány, H a ember, C pedig csirke szekvenciára utal. A konszenzus szekvenciában nagy betűvel jelöltük, ha az összes szekvenciában megegyezik az aminosav. Konszenzus szekvenciát ott irtunk, ahol az adott pozícióban a szekvenciák legalább 51%-a megegyezik. Sötét háttérrel jelöltük a konszenzussal megegyező aminosavakat, világosabb hátérrel a konszenzussal rokon aminosavakat. A rokon aminosav csoportok a következők voltak: I, L, M, V; F, W, Y; D, E; N, Q; H, K, R; S, T; A; C; G; P.

gítségével megszerkesztettük az első hurok dendrogramját is (22. ábra). Látható, hogy először az ős első hurokból kialakult az ős kapcsoló fehérje első hurok, és az ős proteoglükán első hurok. Ezután az ős proteoglükán első hurok modul és megkettőződött, és kialakult az ős aggrecan illetve az ős versican és neurocan első hurok. A dendrogram szerint ezután következett a gerincesek törzsfejlődése, azaz először leváltak a madarak (CLPL1, ill. CAGG1L1) majd az emlősökön belül is megtörtént a fajok szétválása és kialakulása.

Az irodalmi áttekintésben már említettük, hogy az első hurok az Ig modulok családjába tartozik (63). Természetes, hogy megprobáltuk a rendelkezésünkre álló SWISSPROT és PIR fehérje adatbázisokból kikeresni az első hurokkal közös eredetű, Ig családba tartozó egyéb rokon szekvenciákat. A kapcsoló fehérje első hurkának aminosavszekvenciáját mintaként használva kaptuk azt az eredményt, amiből kiderült, hogy



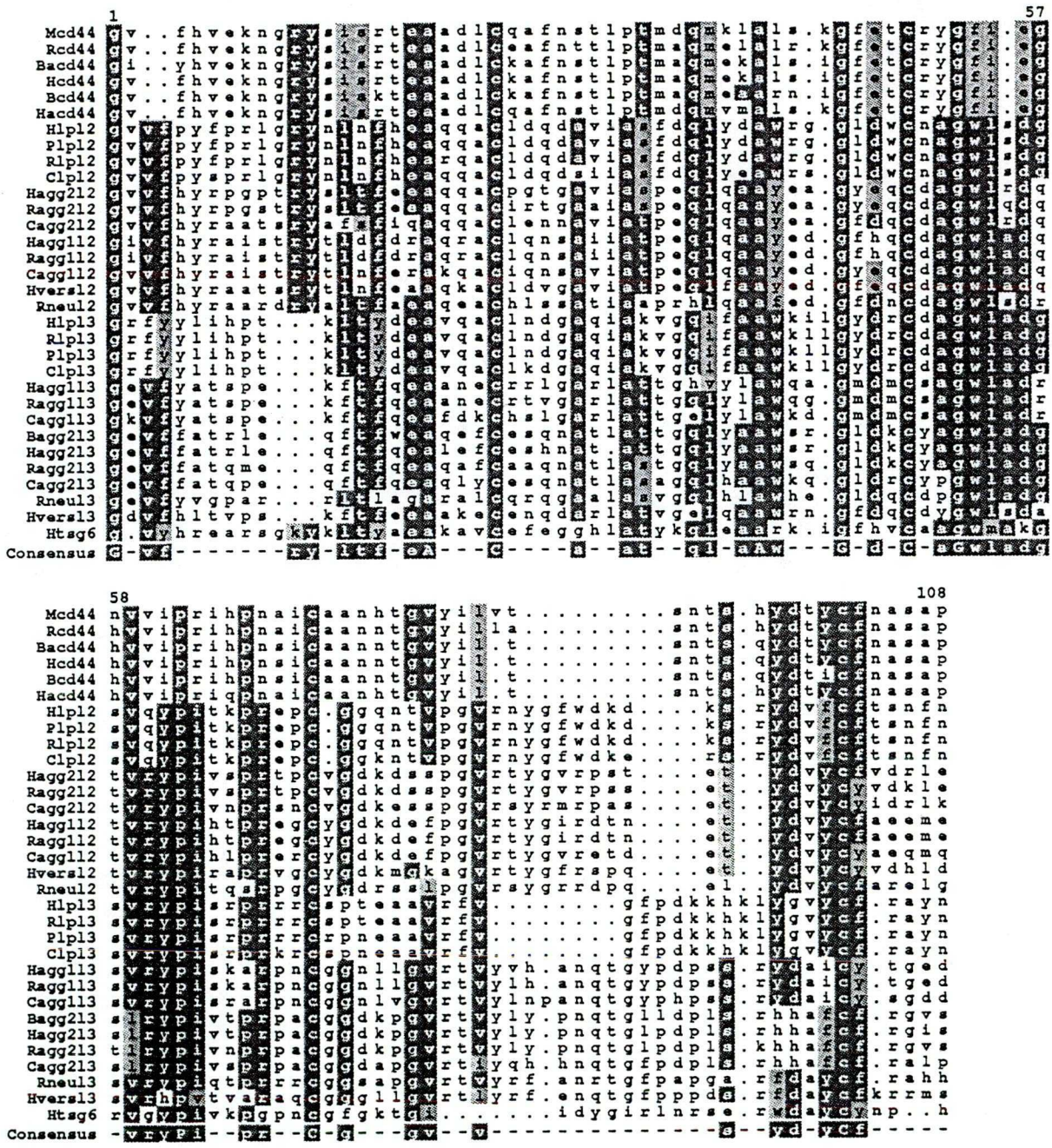
22. ábra: az első hurok szekvenciák rokonságát kifejező dendrogram. A szekvenciák jelölése ugyanúgy történt, mint a 21. ábránál.

olyan nagy a távolság ezek között az egyébként rokon modulok között, hogy a hagyományos kereső programok számára rejtve maradtak a közös vonások. Akkor mégis mi alapján nevezhetjük ezeket a modulokat közös eredetűeknek, hogyan lehet ezt bizonyítani? Egyáltalán melyek a kritériumai a rokonság megállapításának? Az egyik fontos feltétel, mint már az irodalmi áttekintésben is említettük, hogy a modulok azonos fázisú intronokkal legyenek megszakítva. Ez ebben az esetben teljesül, hiszen mind a kapcsoló fehérje első hurok, mind az immunglobulinok variábilis régióiban található Ig modulok egyes fázisú intronokkal vannak megszakítva. Másrészről fontos kritérium, hogy a rokon moduloknál, a szerkezetileg, esetleg funkcionálisan fontos aminosavak és környezetük hasonló pozicióban legyenek, és ezek az aminosavak megegyezzenek egymással, vagy csak olyan cserék legyenek, amelyek nem érintik az adott aminosav karakterét. Így például az egyik

legfontosabb, hogy a harmadlagos strukturát meghatározó, diszulfid kötésben résztvevő ciszteinek pozíciója és száma megegyezzen. Ezért egy olyan módszert alkalmaztunk a távoli rokon modulok megtalálására, a rokonság bizonyítására, amely figyelembe veszi, melyek az összehasonlított első hurok szekvenciák strukturálisan fontos, konzerválódott részei. A már bemutatott (21. ábra) egymáshoz igazított első hurok szekvenciákból a ProfileMake programmal egy úgynevezett profile-t készítettünk. Ez a profile azt tartalmazza, hogy egy adott pozícióban melyik aminosav milyen gyakorisággal fordul elő. Értelemszerűen, ha egy helyen például a cisztein fordul elő az összes szekvenciában akkor a keresés során is csak azok a szekvenciák jöhetnek szóba, ahol szintén cisztein van az adott pozícióban. Az így készült profile alapján a ProfileSearch programmal vizsgáltuk át a SWISSPROT adatbázist. Az eredmény megdöbbentő volt, hiszen több száz Ig-modullal talált hasonlóságot a program, bizonyítva, hogy valóban az első hurok az immunglobulin rokon modulok szuper családjába tartozik.

2. A második hurok vizsgálata

Mint már említettük, a kapcsoló fehérje második és harmadik hurkával, az un. PTR modulokkal rokon modulok szekvenciái rendelkezésünkre állnak a csirke (46), ember (52), sertés (52), és patkány (50) kapcsoló fehérjékből, az emberi (64), csirke (65), marha (101) és a patkány (18) aggrecanból, az emberi versicanból (66), és a patkány neurocanból (67), valamint a ember (76, 77), patkány (80), egér (102), pávián (103), marha (104) és hörcsög (79) cd44 antigénből és ember TSG-6 fehérjéből (78). Összességében a különböző fajokból 32 szekvencia állt rendelkezésünkre, ezeknek az aminósavsorrendjét hasonlítottuk össze egymással a PileUp program segítségével. A kapott eredmény a konszenzus szekvenciával a 23. ábrán látható. A konszenzus szekvenciában nagy betűvel szerepelnek azok az aminosavak, amelyek a legkonzervatívabbak, tehát mind a 32 szekvenciában megegyeznek. A szekvenciák összehasonlítása után felvetődött az a kérdés, hogy lehet-e találni olyan régiót, amely felelős lehet a hialuronsavhoz való kötődésért. Ennek a régiónak nagyon konzervatívnak kell lennie, mivel a hialuronsavhoz való kötődés képessége közös ezekben a modulokban. Ugyanakkor az aggrecan G2 globuláris egységéről kimutatták, hogy nem kötődik a hialuronsavhoz (52), tehát a nyolc darab, a különböző fajok aggrecan G2 globuláris egységeiből származó PTR moduloknak (xAGG2Ly, ahol x a fajnév, y pedig 2 vagy három) ezekben a régiókban különbözni kellene a többi szekvenciától. Az összehasonlított szekvenciákban azonban nem tudtunk egyértelműen olyan régiót kijelölni, amely megfelelne az előbb felállított kritériumnak. Ugyanakkor kísérleti módszerekkel is megpróbálták megta-



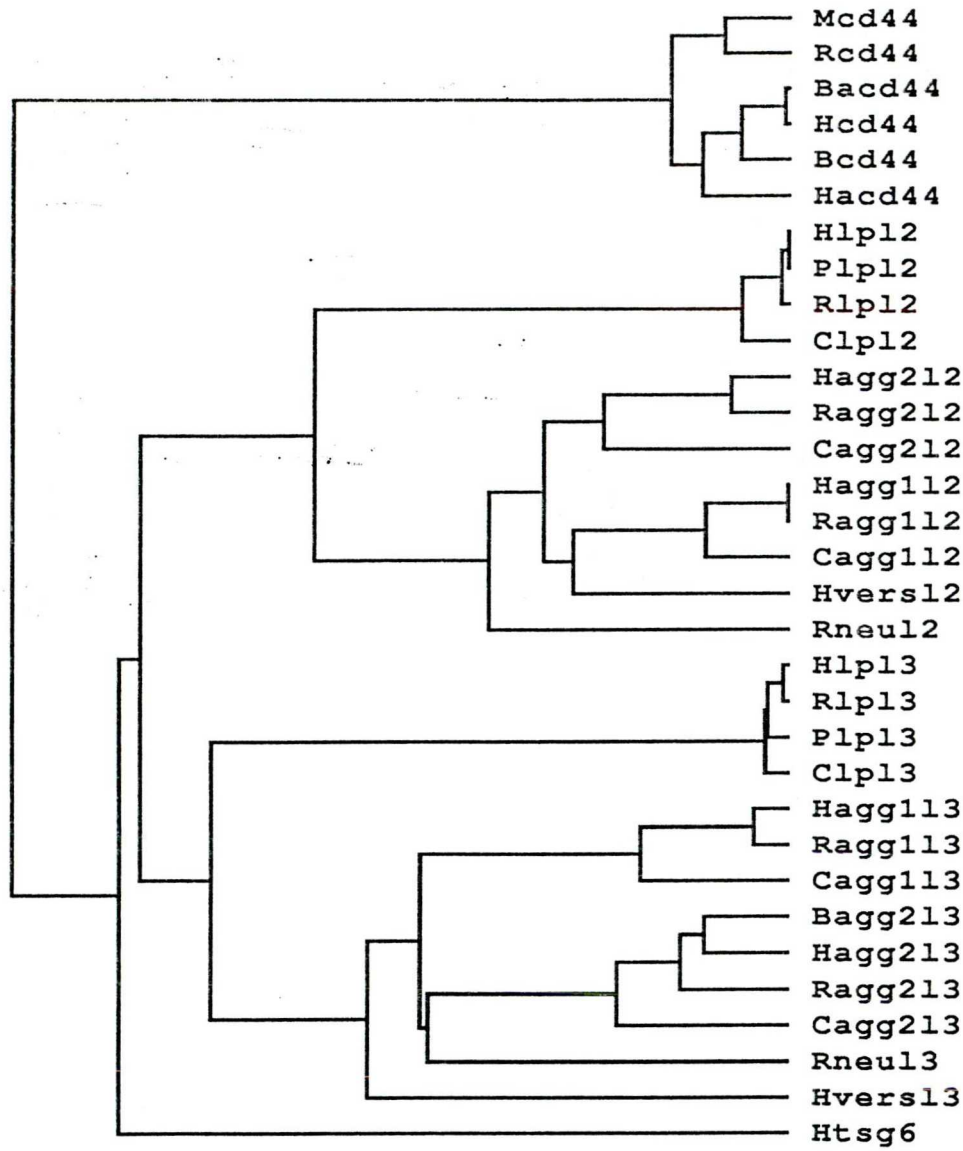
23. ábra: a rendelkezésre álló PTR modul szekvenciák összehasonlítása és a levonható konszenzus szekvencia. A szekvenciák elnevezése hasonló elvek alapján történt, mint a 20. ábránál, azzal a kiegészítéssel, hogy az l2 a duplázodott PTR modulok esetében az N-terminális fele eső modult, az l3 pedig a C-teminális felé esőt jelenti. A g2 a G2 globuláris egységre utal, a cd44 a cd44 antigént, a tsg6 pedig a TSG-6 fehérjét jelenti. A fajneveknél az M az egérre, a Ba a páviánra, a B a marhára, a Ha a hörcsögre utal. A konszenzus szekvencia generálását, és az aminosavak kiemelését a 21. ábránál leirtak szerint végeztük.

lálni, mely régiók lehetnek felelősek a hialuronsavhoz való kötődésért. Goetinck és mti. (105) két ilyen régiót találtak, a 60. pozíciótól a 69.-ig, illetve a 86. pozíciótól a 100-ig terjedőket. Szerintük ezekben a régiókban a pozitívan töltött aminosavak halmozott előfordulása segíti elő a negatív hialuronsav kötődését. Ha megnézzük az összehasonlított PTR modulokat, akkor láthatjuk, hogy valóban az első régióban a CLPL3-ban a 10 aminosavból 5 lizin, hisztidin vagy arginin, mig a CLPL2-ben csak kettő. A különböző cd44 PTR modulokban, amelyek szintén kötnek hialuronsavat, vagy csak egy arginin (Hacd44), vagy egy arginin és egy hisztidin található. A TSG-6 tumor nekrózis faktor indukált fehérjében is csak egy pozitív aminosav (lizin) van ebben a régióban. Ugyanakkor az aggrecan G2 globuláris egységében található PTR modulokban 2 pozitív aminosav esik erre a szakaszra, pedig ezek a modulok nem kötődnek a hialuronsavhoz. Hasonló a helyzet a másik régiónál; 3-3 pozitívan töltött aminosav található a kapcsoló fehérjék harmadik hurkánál vagy az aggrecan G2 globuláris egységében levő második PTR modulnál (pgg2l3), de egy sem található például a cd44 fehérjéknél, a proteoglükánok két globuláris egységének második PTR moduljainál, vagy a TSG-6 fehérjénél.

Perkins és mti. (52) szintén az első régióba térképezik a hialuronsavhoz való kötődésben létfontosságú szerepet játszó aminosavakat. Ők azt mondják, hogy az ötödik pozíció nagyon fontos, mivel itt az aggrecan G2 egységében levő PTR modulokban valin található szemben a többi PTR modullal, melyekben ebben a pozicióban treonin, szerin, arginin vagy hisztidin fordul elő. Megjegyzendő azonban, hogy az összehasonlításunk szerint ugyancsak valin található ebben a pozicióban a TSG-6 proteinben is, a cd44 fehérjékben pedig izoleucin, ami nagyon hasonlít a valinra. Összességében megállapíthatjuk, hogy a szekvenciák összehasonlítása alapján nem lehet egyértelműen egy-egy régiót, vagy egy-egy aminosavat hozzárendelni a hialuronsavhoz való kötődés képességéhez.

3. A kapcsoló fehérje molekuláris evolúciója

Az összehasonlítások alapján felrajzolt két dendrogramból (23., 24. ábra) világosan látszik, hogy a PTR modul molekuláris evolúciója időben megelőzte a gerincesek törzsfejlődését, azaz előbb alakultak ki a genomon belül a modul különböző fehérjékben szereplő formái, mint a madarak és az emlősök szétválása. A 25. ábrán ennek a genomon belüli molekuláris evolúciónak a feltételezett lépései láthatóak. Mivel a legtöbb PTR modul az emberből ismert, ezért az emberi PTR modul szekvenciák összehasonlítását vettük alapul. Kiindulásképpen egy ős PTR modult és egy ős első hurok Ig modult tüntettünk fel. Megjegyzendő, hogy az itt szereplő Ig modul már feltételezhetően átment egy molekuláris

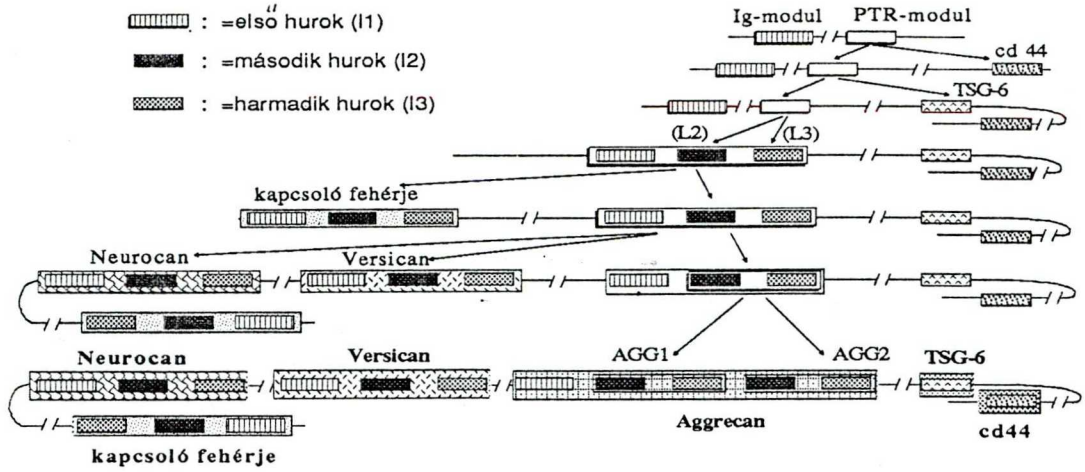


24. ábra: a PTR domének rokonságát kifejező dendrogram. A szekvenciák jelölése az előzőekben ismertetett módon történt.

evolúción, amely során a többi Ig szupercsaládba tartozó modul kialakulhatott. Azt se lehet teljességgel kizárni, hogy az ábrán feltüntetett ős PTR modul maga is korábbi evolúció erdménye és más rokon modulokkal közös a származása.

Tehát ennek az ősi PTR modulnak az első megkétszereződésekor az egyik termékből kialakultak a cd44 fehérjék, míg a másik később mégegyszer megduplázódott. Ekkor vált le a TSG-6 fehérjében található PTR modul. Ezután következett a későbbi kapcsoló fehérje és a proteoglükánok második és harmadik hurkának a kettéosztódása. Az evolúció ennek a

pontjának közelében játszódhatott le az a folyamat, melynek során az ős első hurok Ig modul és az ős második és harmadik hurok modulok egymás közelébe sodródhattak, és a következő duplikációt, melynek során az ős kapcsoló fehérje kialakult, már együtt szenved-



25. ábra: a mai napig ismerté vált, és PTR modult tartalmazó fehérjék exon shuffling által történő kialakulásának a modellje.

ték el. A másik termék újabb osztódása során keletkezett a versicanban lévő globu-láris egység, és az ős aggrecan globuláris egység is. Ezidőtájt alakulhatott ki a versicannal párhuzamosan a neurocan is, pontos becslést azonban nem tudunk adni, mivel a mai napig csak a patkány neurocan szekvenciája ismert, és a neurocan-versican elválás időben nagyon közel egymáshoz történhetett. Érdekes, hogy az ezt követő lépésben az aggrecan úgy osztódott, hogy kimaradt az első hurok, tehát igy jöhetett létre az aggrecan csonka második globuláris egysége. Összeségében tehát hat duplikációra volt szükség ahhoz, hogy létrejöjjön az emberben jelenleg ismert tíz darab PTR modul.

A gerincesek törzsfejlődése során az első hurok és a PTR modul dendrogramjai alapján, összhangban az evolúció kutatás eredményeivel, először a madarak (Clp..) és az emlősök váltak el, majd az emlősök között először a rágcsálók váltak le (patkány, egér hörcsög), majd a marha és a sertés, és a vizsgált fajok között a Cd44 szekvenciák tanusága szerint az emberhez a pávián áll a legközelebb.

EREDMÉNYEK MEGVITATÁSA

I. A kapcsoló fehérje gén 5' vége

Specifikus oligonukleotid primerek segítségével, amelyek a kapcsoló fehérje gén második és harmadik exonjainak az 5' végéhez hibridizálnak, 5' végi specifikus cDNS génkönyvtárat készítettünk. Erre két okból volt szükség:

1. A primer extenziós kísérletek felfedték, hogy a génnek 500 bp-nál hosszabb 5' végi nem transzlálódó vége van, amely jócskán túlnyúlik a gén eddig ismert 5' végi részén. Mivel ezekben a kísérletekben többféle hosszúságú kapcsoló fehérje specifikus termékeket kaptunk, valószínűnek tűnt, hogy a génnek erről az 5' végi részéről átírt mRNS valóban többféle hosszúságú, amelynek oka lehet többszörös iniciáció, vagy a pre-mRNS-ből alternatívan kivágódó különböző méretű intronok létezése.
2. A gén intronjaiban található repetitív szekvenciák lehetetlenné tették, hogy az első exont és a promoter régiót hordozó genomiális szakaszt átfedő kromoszómális fragmentumok segítségével izolálni lehessen. Ezért szükségessé vált, lehetőleg a teljes első exon klónozása, hogy próbaként használva izolálni lehessen az említett genomiális régiókat.

Az kapott, 5' végre specifikus cDNS klónkönyvtárból 2000 telepet vizsgáltunk meg telephibridizációval és 32 pozitívat találtunk, ami több, mint a klónok másfél százaléka. A többi klón keletkezésére az adhat magyarázatot, hogy a klónozás első lépése, a primer extenzió során nagy szerepet játszik az ún. "self-priming", amikor az RNS felhajló vége, vagy egy másik odahibridizált RNS szolgál primerként. Később, az így keletkezett DNS ugyanúgy klónozódik, mint a "normális", az általunk hozáadott primerek meghosszabbításával nyert. Meghatároztuk a kiválasztott klónok inszertjeinek méreteit, restrikciós térképeiket, valamint húsz klónnak a szekvenciáját is. Kiderült, hogy a 32 klón fele nem kapcsoló fehérje eredetű. Ezt bizonyítják egyrészt a szekvencia adatok, hiszen ezeknél a klónoknál az első húsz nukleotid teljesen megegyezik a hozzáadott primer, így a kapcsoló fehérje gén megfelelő részének a szekvenciájával, majd egészen más, idegen szekvencia következik. Másrészt a Southern hibridizációs kísérletek eredményeiből (9. ábra) is látszik, hogy ezek a klónok nem hibridizálnak a kapcsoló fehérje eredetű próbával. Egy klónt próbaként használva Northern hibridizációt is végeztünk. Az eredmény egyértelműen bizonyítja, hogy az adott klón, a p8A12, nem kapcsoló fehérje eredetű, mivel az RNS amihez hibridizál, nem porcspecifikus

(csirke koponyatető csontban is előfordul), és más méretű mint a kapcsoló fehérje mRNS. Az ilyen klónok előfordulását magyarázza, hogy hiába hibridizálódik a primer olyan helyre, amelynek a szekvenciája nem komplementer teljesen vele, a második szál szintézisekor a primer szolgál mintaként, és így nem az eredeti szekvencia tükröződik az ilyen klónokban. Meghatároztuk a leghosszabb cDNS klón, a pLP7G12 szekvenciáját is. A klón 5' végét próbaként használva Northern hibridizációval bizonyítottuk, hogy az teljes hosszúságában kapcsoló fehérje mRNS eredetű. Ennek a klónnak az első exont tartalmazó részét használtuk fel próbaként csirke genomiális könyvtár átvizsgálására. Egy pozitív klónt találtunk, amely 21 kb-os genomiális DNS-t hordoz. Az inszert végeit, és a *Eco*RI, *Hin*dIII fragmentumait szubklónoztuk . Southern hibridizációk segítségével meghatároztuk a fizikai térképét (13. ábra). Az első exont hordozó szakaszt a klón jobb kar felé eső végén lokalizáltuk. Ez egyben azt is jelenti, hogy a klón az első intronból is hordoz mintegy húsz kb-nyi részt. A korábbi vizsgálatokból ismert λgLP12.1 klón (54), amely a második exont hordozza, ugyancsak tartalmaz az első intronból mintegy 14 kb-t, amely nem mutat kereszthibridizációt az első intronnak a λgLP100 klónban található részével, és a restrikciós térképeik is eltérnek egymástól. A két klón tehát az első intronból 33 kb-t hordoz, az intron mérete azonban ennél is hosszabb. Összességében a kapcsoló fehérje génből az eddig izolált genomiális klónok mintegy 80 kb-t fednek le, de a gén hossza a genomban nagyobb, mint 100 kb, tehát az elsődleges transzkriptum is ilyen hosszú!

A λgLP100 klónból szubklónoztuk az első exont hordozó 1.7 kb-os *Eco*RI-*Sal*I fragmentumot, és meghatároztuk a szekvenciáját. Primer extenzióval és T4 polimerázzal végzett térképezéssel meghatároztuk a kapcsoló fehérje gén transzkripciós kezdőpontjait. Ahogy a 19. ábrán látszik, nem egy, hanem több startpont is létezik független kísérletek egybehangzó eredményei szerint. Ezek a pontok a -532 - -482 DNS szakaszon találhatóak, ahol egy hosszabb pirimidin traktus található. Érdekesség, hogy a főbb iniciációs pontok a traktusban található timineket kb. tíz bázisonként megszakító citozin bázisoknál találhatóak. Az iniciációs helyek közül leggyakrabban használt az 5' végtől számított harmadik csoport, ahol a 8G1 típusú transzkriptumok nagy része, és valószínűleg maga az izolált pLP8G1 klón is kezdődik. Ezt úgy bizonyítottuk, hogy olyan specifikus oligonukleotid primerrel végeztük a primer extenziót, amelyiknek a közepére esik a kivágódási hely, így csak a 8G1 típusú RNS-hez hibridizálhat teljes hosszúságban. Ennek tükrében meg kell jegyezni, hogy a pLP8G1 klón egy bázispár híján valóban teljes hosszúságú cDNS klón, ami bizonyítja cDNS klónozási stratégiánk helyességét.

Az iniciációs pontoktól 30-70 bp-nyira található egy TATA box szerű szekvenciamotívum, kissé távolabb pedig két CCAAT box motívum. Jellemző még a promoter régió szekvenciájára, hogy tartalmaz egy-egy 47 és 32 bp hosszú purin bázisokból álló traktust és egy 46 bp-ból álló pirimidin traktust is. Ezek szerepe egyelőre nem tisztázott. Meghatározták az emberi és a patkány kapcsoló fehérje gén 5' végét is (106). Patkány kondroszarkóma kapcsoló fehérjét kódoló cDNS-ek közül a leghosszabbak is csak 15 bp-t tartalmaztak a gén első exonjából, ezért Rhodes és munkatársai hasonlóan az e dolgozatban leirtakhoz, 5' specifikus cDNS könyvtárat készítettek kapcsoló fehérje specifikus oligonukleotid primer segítségével. Megvizsgálták a pozitív cDNS klónokat, és azt találták, hogy a leghosszabbak is csak 20-25 bp-ral nyúltak tovább 5' irányban az addig ismert szekvenciánál. Ezzel együtt az első exonból 40 bp állt rendelkezésükre, ez alapján szintetizáltak egy oligonukleotid próbát, amellyel patkány genomiális könyvtárat vizsgáltak át. Egy pozitív klónt találtak, amelyből az első exont tartalmazó résznek meghatározták a szekvenciáját. Oligonukleotid próbával sikerült izolálni 3 átfedő, az emberi kapcsoló fehérje gén első exonját hordozó genomiális klónt is. Primer extenzióval és S1 nukleázos térképezéssel sikerült meghatározni mind az emberi, mind a patkány gén transzkripciós kezdőpontjait. Eredményeik szerint a patkány kapcsoló fehérje gén első exonja mindössze 62 bp-ból áll, míg az emberi gén első exonja ennél 29 bp-ral hosszabb. Ugyanakkor Dudhia és Hardingham (53) emberi ízületi porcból olyan kapcsoló fehérje cDNS klónt izolált, amely 208 bp-t tartalmaz az első exonból, tehát a Rhodes és mti. által meghatározott ember első exonnál eleve 107 bp-ral hosszabb. A túlnyúló rész szekvenciája 100%-ban homológ a Rhodes és mti. által már promoter régióként jellemzett szekvenciával. Tehát az emberi kapcsoló fehérje génnél minimum két promoternek kell léteznie. Hasonlóképpen (mivel a csirke és a ember, mind a csirke és a patkány kapcsoló fehérje gén 5' nemtranszlált régiói között kb. 80%-os homológia van) a csirke első exon is túlnyúlik a Rhodes és mti. által embernél és patkánynál megállapított első exon 5' végénél. Ráadásul mind a patkány-ember, mind a patkány-csirke viszonylatában olyan konzervatívak a szekvenciák, hogy mind a patkány gén transzkripciójának kezdőpontja, mind a patkány promoter TATA box motívuma bázisra megegyeznek a csirke és az ember szekvenciájával. Ezzel kapcsolatban a következő kérdések vetődnek fel:

- Miért van az ember és a patkány gén kezdőpontja között 29 bp különbség amikor a szekvenciák szinte teljesen azonosak?
- Hol van az emberi gén második promotere, amiről a Dudhia által izolált cDNS íródik át?

- Létezik-e a patkánynál is a távolabbi promoter?
- Létezik-e a csirkénél a közelebbi promoter, és ha igen, melyik kezdőpont érvényesül, a patkányra, vagy az emberre jellemző?

Jelenlegi tudásunk alapján ezekre a kérdésekre még nem tudunk pontos választ adni. Annyi azonban bizonyos, hogy a kiindulásként használt mRNS patkány esetében kondroszarkoma sejtekből, a ember promoter kimutatásakor magzati porcból, míg Dudhia által izolált továbbnyúló emberi kapcsoló fehérje cDNS klón esetében felnőtt ízületi porcból, a csirke esetében pedig 14-15 napos embriókból származik. Nincs kizárva tehát, hogy alternatív promoter használat történik. Például a csirke embriók végtagkezdeményeiben a porcos differenciálódás megindulásakor a közeli promoter aktív, majd később a 14 napos csirke embriókban, vagy a már kialakult ízületi porcban a távolabbi promoterről íródnak át a kapcsoló fehérje hírvívő RNS-ek. Természetesen az sem zárható ki egyértelműen, legalábbis a primer extenziós kísérleteink alapján, hogy a 14 napos csirke embriók porcos mellcsont kezdeményében mind a kétfajta promoter aktív, vagy hogy a csirke esetében csak a távolabbi promoter hasznosul. Ennek eldöntése a későbbi vizsgálatok feladata lesz.

A patkány kondroszarkoma kapcsoló fehérje génnél kimutattak a transzkriptumok egy tizedénél egy alternatív exont a második és harmadik exonok között (55). Az ezek által a hírvívők által kódolt fehérje 53 aminosavval többet tartalmaz, funkciója ismeretlen. Megnéztük, hogy a csirkénél kimutatható-e ez az alternatív exon. Vizsgáltuk egyrészt az izolált, a prE3 primerről indult cDNS klónokat, másrészt PCR-t végeztünk az első és a harmadik exonok között. A feltételezett alternatív exont egyik módszerrel sem tudtuk kimutatni, tehát valószínűleg ebben is különbözik egymástól az általunk vizsgált 14-17 napos csirkeembriók mellcsont kezdeményéből, és a patkány kondroszarkómából származó kapcsoló fehérje gén.

A második exont tartalmazó 11/8-as próbával pozitívnak bizonyult 11 klón közül tíznek határoztuk meg a szekvenciáját, és azt találtuk, hogy mind a tíz megegyezik a genomiális rész nukleotidsorrendjével. Négy klónnál azonban a cDNS szekvenciája a genomiálishoz képest kisebb-nagyobb deléciókat tartalmaz. Mivel a deléciókat konszenzus kivágódási akceptor és donor helyek határolják (3. táblázat), arra a következtetésre jutottunk, hogy alternatív intronhasználat történik a kapcsoló fehérje gén 5' végén. Azt, hogy ezek a cDNS-ek nem klónozási műtermékek, úgy bizonyítottuk, hogy az érintett régióról átírt cDNS-eket PCR segítségével felszaporítottuk. Ha csak egyfajta, az intron nélküli cDNS létezne, akkor a PCR során is csak ez a termék szaporodna fel. Ezzel szemben a PCR

eredményeképp olyan fragmentumokat kaptunk, amelyek megfelelnek a cDNS-ek típusai által számítottaknak. Ezenkívül olyan termékeket is kaptunk (Hyp1-Hyp4), amelyeket méretük allapján le lehet vezetni a már az ismertetett cDNS klónoknál hasznosult kivágódási donor és akceptor helyek újfajta kombinációival. Ezeknél a feltételezett mRNS típusoknál tovább nő a kihasadt intronok együttes hossza, így az 5' végi nem transzlált régió is lerövidül. Egyedül a teljes hosszúságú, tehát az intron nélküli cDNS klónnak megfelelő fragmentumot nem sikerült kimutatni. Ennek alapján felmerül a kérdés, hogy ez a pLP7G12 klónnak megfelelő kivágódás nélküli típus nem egyszerűen egy köztes, kivágódások előtt álló pre-mRNS-e (107)? Ez ellen két fontos érv szól: Az egyik, hogy az izolált, a nem transzlált régióban teljes hosszúságú cDNS klón, a pLP7G12, tartalmazza az első exonon kivül a másodikat teljes hosszúságban, valamint a harmadikból is egy rövid darabot, tehát olyan RNS-ről iródott át, amelyikből már kivágódott az első és a második intron. A másik érv az az, hogy a pLP7G12 klón 5' végi *Ava*II helyétől az 5' végéig húzódó fragmentummal Northern hibridizációt végeztünk, és sikerült ezzel a próbával is kimutatni mind a három, a kapcsoló fehérjére jellemző mRNS-t (11. ábra), pedig ez a fragmentum olyan RNS darabot reprezentál, amelyik az 5S3, 8C7, és 8G1 típusú RNS-ekből teljes egészében, a 11B4 típusú RNS-ből pedig részlegesen kivágódik.

Ezek az eredmények bizonyítják, hogy a csirke kapcsoló fehérje gén 5' végén alternatív splicing történik. Az alternatív splicing általánosan előforduló jelenség a magasabbrendű szervezetekben. Le kell azonban szögezni, hogy az esetek nagy részében ez a folyamat fehérjét kódoló szakaszokat érint, ami azt jelenti, hogy a különböző típusú RNS-k különböző fehérjetermékeket kódolnak, melyeknek más és más lehet a funkcójuk. Ugyanakkor a kapcsoló fehérje esetében pontosan ugyanaz a fehérje íródik át az összes típusú RNS-ről. Hasonlóképpen, az 5' nem transzlálódó régióban mutattak ki alternatív splicing-ot a patkány I típusú inzulin-szerű növekedési faktor (108), a patkány acetil-koenzim A karboxiláz (109, 110), és a sertés muszkarin típusú acetilkolin receptor génnél (111). Természetesen felvetődik a kérdés, hogy mi lehet a jelentősége, vagy egyáltalán miért következik be a kódoló részt nem érintő alternatív splicing? Az egyik lehetséges magyarázat szerint ezt a régiót semmilyen evolúciós nyomás nem befolyásolja, mivel nincs megváltozott fehérjetermék, tehát az itt lévő kivágódási helyek véletlenszerűen hasznosulnak. E mellett szól az, hogy, amint az a 3. táblázatból leolvasható, a hét lehetséges donor helyből öt, a négy lehetséges akceptor helyből pedig kettő használatára találtunk példát a cDNS-ek között. Deák F. és mti. PCR termékek szekvenciájának a vizsgálatával a fennmaradt kettő donor helyből az egyik használatára találtak példát (112). A véletlen szerepét bizonyítja a pLP8C7, a

pLP5S3 és a pLPD48 (112) klónok megléte is, hiszen ezeknél a klónoknál a kivágódott intron hosszában mindössze 4-4 nukleotid különbség van, amit aligha lehet mással magyarázni, mint a három egymást átfedő donor hely véletlenszerű használatával.

A másik lehetséges magyarázat szerint a különböző RNS formák különböző transzlációs aktivítással birnak azáltal, hogy befolyásolják a transzláció iniciációját. Jelenlegi ismereteink szerint ez ötféleképpen történhet (107). Megvizsgáltuk, hogy a kapcsoló fehérje különböző mRNS formái esetében milyen elvi lehetőségei lehetnek a transzlációs szintű szabályozásnak:

- Az iniciációs AUG kodon és közvetlen környezetének szekvenciája hatással lehet a transzláció megkezdődésére. Mind a kísérletek során (113-115), mind 699 gerinces mRNS összehasonlításával (100) ugyanazt a konszenzus szekvenciát találták ($GCC^{A}/_{G}CC\underline{AUG}G$). Ez azt jelenti, hogy egyrészt a kísérletek tanúsága szerint ennél a szekvenciánál a legerősebb a transzláció, másrészt ezek a nukleotidok fordulnak elő legtöbbször az adott pozicióban az összehasonlított mRNS-eknél. A kapcsoló fehérje esetében az iniciációs AUG kodon környezete (GTGAA**AUG**A) mindössze két pontban egyezik meg a konszenzussal, tehát gyengének nevezhető. Ez azt is jelentheti, hogy nagyobb szerepe lehet az egyéb transzlációs szabályozó mechanizmusoknak.
- Lényeges lehet a transzláció iniciálásában, hogy létezik- vagy léteznek-e a tényleges AUG kodonoktól 5′ irányban nyitott leovasási keretek. A gerinces gének 90%-ánál a transzláció iniciációja az első AUG kodonnál történik (100). A fennmaradó 10%-nál a "ráadás" AUG kodonok hatással lehetnek a transzláció iniciációjára. A vizsgálatok szerint ez a hatás gátlás, tehát a ráadás AUG kodonok kiejtése megemeli a transzláció hatásosságát. A kapcsoló fehérje esetében kettő ilyen plusz AUG kodon található az 5′ nem transzlált régióban a 7G12 típusú RNS molekulában. Ezek közül az egyik egy 42 aminosavból álló nyitott leolvasási keretet határoz meg. Arra azonban semmiféle indikáció sincs, hogy ez a rövid peptid in vivo termelődik-e vagy nem. Ráadásul mind a két kodon környezete teljesen eltér az előzőekben ismertetett "ideális" konszenzus szekvenciától. Meg kell jegyezni, hogy a 7G12 és a 11B4 típusú RNS-ek kivételével ez a két AUG kódon is kivágódik az alternatív splicing során.
- Értelemszerűen befolyásolja a transzláció iniciációját az 5′ nem transzlált régió hossza, hiszen a "vándorló" model alapján a riboszoma először az mRNS 5′ végéhez kötődik, majd onnan 3′ irányba vándorolva keresi meg az iniciációs kodont (116). A teljes hosszúságú 7G12 típusú mRNS mintegy 500 nukleotid 5′ nem transzlálódó

részt hordoz. Ez nem általános, hiszen a gerinces gének alig több mint 1%-ának van 500 bp-nál hosszabb 5' nem transzlálódó régiója (117). A vizsgált mRNS-ek 75%-ánál ez 100 nukleotidnál rövidebb. (100). Ugyanakkor a többi típusú mRNS-ekből az alternatív splicing miatt ez a régió rövidebb, tehát a különböző típusoknál más- és más lehet az iniciáció sebessége.

- A kísérletek szerint hatással van a transzláció iniciációjára az RNS-nek az 5' nem transzlálódó régióban és az AUG kodon környezetében kialakuló térszerkezete. Az itt előforduló stabil hajtűhurkok specifikus kötőhelyként szolgálhatnak bizonyos fehérjék számára, amelyek aztán befolyásolhatják a transzlációt. A ferritin mRNS esetében az 5' véghez kötődő fehérje alapállapotban gátolja a transzlációt. Vas adására ez a gátlás oldódik, és megindul az aktiv transzláció (118). Az 5' nem transzlált régióban kialakuló RNS hajtű hurkok önmagukban is képesek gátolni a transzláció megkezdését azáltal, hogy megakadályozzák a stabil riboszomális 40S alegység- mRNS komplex kialakulását (119), vagy éppen a riboszómális alegység 3' irányba történő vándorlását (120). Az AUG kodontól 3' irányban kialakuló nem túl stabil RNS hurok elősegítheti a transzláció megindulását egy "gyenge" iniciációs kódon esetében (121). Az ilyen hajtű ugyanis csak éppen annyira "lassítja" le a riboszóma vándorlását, hogy az biztosan "észrevegye" az AUG kodont. Megvizsgálva a kapcsoló fehérje 5' végét azt találtuk, hogy az erős másodlagos szerkezet kialakítására képes. A különböző típusú RNS-ek más és más hurkokat képeznek. A legstabilabb másodlagos szerkezet a 7G12 típusú RNS esetében jósolható, míg a legkevésbé stabil a 8G1 típus esetében, amely a vizsgálataink szerint a leggyakoribb. Érdekes ugyanakkor, hogy mindkét típus esetében megtalálható a 21. ábrán B-vel jelzett hajtű, ami viszont hiányzik például az 5S3 típusnál. Vizsgálódásaink természetesen magukba foglalták az AUG kodontól 3' irányba eső részt is, de nem találtunk olyan hajtű hurkot, amely erősíthetné a transzláció iniciációját.

Az előzőekben leírtak alapján megállapíthatjuk, hogy a kapcsoló fehérje 5' végét érintő alternatív splicing során keletkezett különböző típusú mRNS-ek különböző transzlációs aktivítással bírhatnak. Természetesen ennek akkor lehet csak igazán jelentősége, ha feltételezzük, hogy az eltérő mRNS variánsok eltérő szövetekben, vagy sejttípusokban expresszálódnak, vagy ugyanazon szövet illetve sejttípusban az egyedfejlődés során egyszer megjelennek, egyszer eltűnnek. A helyzetet nehezíti, hogy nyilvánvalóan az általunk vizsgált kiindulási anyagban, a 14 napos csirke embriók porcos mellcsont kezdeményében, valamennyi forma megtalálható, bár kétségkívül nem azonos arányban, hiszen a 8G1 típusú

mRNS mutatkozott a legnagyobb mennyiségűnek. Ugyanakkor a csoport más munkatársainak vizsgálatai szerint a juvenilis csirkék mellcsontjából illetve ízületi porcából származó porcsejtjeiben már hiányoznak egyes alternatív mRNS formák.

II. A kapcsoló fehérje evolúciója

A kapcsoló fehérje a lehasadó szignálpeptiden kivül három különálló globuláris doménből, hurokból áll. Az N-terminális vég felől az első hurok, amelyet egy diszulfid kötés stabilizál az immunglobulinok variabilis régióival mutat homológiát (63). Hasonló domén található még az összes, eddig ismert aggregálódó proteoglükánban (18, 64, 65), a versicanban (66) és a neurocanban (67) is. A kapcsoló fehérjében és a proteoglükánokban található első hurok domének azonban olyan távoli rokonai az Ig-szuperfamiliába tartozó egyéb doméneknek, hogy a hagyományos kereső, és összehasonlító programokkal ki sem mutatható, mivel a hasonlóság csak egy-két, a másodlagos és harmadlagos szerkezet kialakításában fontos szerepet játszó aminosavra korlátozódik. Ezt bizonyítandó, egy olyan programmal is elvégeztük az analízist, amely fokozottan súlyozza azokat a pozíciókat, ahol a mintaként használt kapcsoló fehérje és proteoglükán szekvenciákban egy-egy aminosav konzerválódott. Az eredmény azt bizonyította, hogy valóban, a teljesen eltérő funkció ellenére ezek a domének strukturálisan megegyeznek, egy közös ősre vezethetőek vissza. A rokonság azonban nagyon távoli, az ős Ig-domén feltételezett duplikációja, melynek során a "kapcsoló fehérje" ág levált, időben valószínüleg sokkal régebben következett be, mint a kapcsoló fehérje, és a proteoglükánok elválása egymástól, mivel ezek sokkal jobban hasonlítanak egymáshoz, mint az Ig-szuperfamilia más tagjaihoz.

Elkészítettük az ismert első hurok szekvenciák összehasonlítását (22. ábra). A szekvenciák egymás közti rokonságát dendrogram rajzolásával illusztráltuk (23.ábra).
A kapcsoló fehérje másik két hurka az ún. PTR modulok közé tartozik (52). Ma összesen 32 ilyen szekvencia ismert. Összehasonlításukkal többek közt szerkezetbeli magyarázatot próbáltunk keresni arra a kísérletekkel igazolt ellentmondásra, hogy az összes többi PTR-doménnel ellentétben az aggrecan G2 globuláris egységében lévő PTR domének miért nem kötődnek a hialuronsavhoz. Amennyiben egy-egy aminosavnak. vagy egy rövidebb régiónak van ebben kitüntetett szerepe, akkor a szekvenciák vizsgálatakor ennek ki kell derülnie. A várakozással ellentétben nem sikerült ilyen különbözőséget találni. Nem tudtuk alátámasztani Goetinck és mti. hipotézisét (105), miszerint az általuk talált két régióban lévő pozitív töltést hordozó aminosavak lennének felelősek a kötődésért, mivel nem lehet

konzekvensen azt mondani, hogy a G2 egység PTR modulokban kevesebb ilyen aminosav van, mint a többiben, hiszen például a cd44 antigén szekvenciákban még kevesebb van, pedig ezek is kötődnek a hialuronsavhoz. Hasonlóképpen lehet cáfolni Perkins és mti. hipotézisét is (52), persze csak ha bizonyítottnak vesszük, hogy a G2 PTR modulok valóban nem kötődnek a hialuronsavhoz. Hogy lehet akkor magyarázni a hialuronsavhoz való kötődés képességét, és egyáltalán mi lehet akkor a G2 globuláris egység szerepe az aggrecanban? Az első kérdésre a választ valószínüleg a variabilis régiókban lévő aminosavak együttes hatásában kell keresni sokkal inkább, mint egy-egy aminosavhoz, vagy rövidebb régióhoz hozzárendelni a hialuronsavhoz kötődés képességét. A második kérdésre egyes kutatók azt válaszolják, hogy a G2 egységnek pusztán térkitöltő szerepe van (52). Ezt igazolja legalábbis, hogy szerkezet kialakításában fontos szerepet játszó, konzervatív, a konszenzusban is megjelenő aminosavak a hialuronsavat nem kötő modulokban is azonosak a többiekével. Ugyanakkor azt várnánk, hogy a többi pozícióban sokkal jobban eltérjenek a hialuronsavhoz kötődő moduloktól, hiszen ezekre a helyekre nem nehezedik evolúciós nyomás. Ez azonban nincs így, a G2 globuláris egységben található PTR modulok csak annyira különböznek a többitől, mint azt a normális, tehát a hialuronsav kötő funkció irányából is ható evolúciós nyomás esetén feltételezhető. Az is elképzelhető, hogy a G2 egység a hialuronsavhoz való kötődés képességét evolúciós léptékkel nézve, csak a közelmúltban veszíthette el, úgyhogy nem volt "ideje" igazából megváltozni azóta.

Az első hurok és a PTR modulok dendrogramjai alapján világosan látszik, hogy a különböző mozaik fehérjék kialakulása, tehát az exon shuffling, megelőzte a gerinces makroevolúciót, tehát a különböző fajok kialakulását. Ez logikus is mivel valószínüleg a porcos halaknál is ugyanúgy megtalálható a porcban a kapcsoló fehérje és az aggrecan mint az embernél. A 23. és a 25. ábrákon látható dendrogramok tehát kétféle folyamatról adnak információt. Egyrészt egy genomon belüli, ha úgy tetszik szomatikus, egyeden belüli, vagy horizontális eseménysorozatról, az exon shufflingról, melynek során DNS illetve kromoszóma darabok kettőződnek meg és rendeződnek át új fehérjéket kialakítva a különböző exonok egymás mellé rendeződésével, másrészt egy makroevolúciós, ha úgy tetszik vertikális folyamatról, melynek során kialakulnak a gerinces fajok. Ha tehát az exon shufflingot akarjuk modellezni, akkor elég csak az egy fajon belül lévő modulokat vizsgálni. Mivel az emberbél ismert jelenleg a legtöbb PTR modul, mi is ezek alapján állítottuk fel modellünket (26. ábra). Kiindulásként egy ős PTR és egy ős első hurok modult feltételeztünk, amelyeknek persze lehetnek még korábbi elődeik is, mint azt az első hurok és az Ig-domének esetében említettük is. Az embernél a jelenleg ismert ötféle PTR modult is

tartalmazó fehérje létrejöttét összesen hat duplikációs eseménnyel tudtuk leírni. Az első hurok exonja valószínűleg az ős PTR domén harmadik megkettőződése után sodródhatott a már párosan álló PTR modulok elé, és a következő duplikációt melynek során az ős kapcsoló fehérje alakult ki, már együtt szenvedhették el. Érdekes ugyanakkor, hogy a modellünkben az utolsó duplikációból kimaradt az első hurok, így alakulhatott ki az aggrecan két globuláris egysége. Az ábrán szereplő neurocan (67) szekvencia a patkányból származik, az emberből még nem izolálták. A szekvencia összehasonlításból és a dendrogrammokból úgy tűnik, hogy a versicannal egy időben alakulhatott ki, erről pontosabb képet csak akkor kaphatunk, ha majd rendelkezésünkre állnak a ember neurocan modulok szekvenciái is.

A különböző fajokból ismert, azonos modulok vizsgálata jól modellezi a fajok evolúcióját, az adatok interpretálásával azonban óvatosan kell bánni, mivel például a génkonverziós jelenségek megváltoztathatják az eredményt. Ezt bizonyítja, hogy míg a kapcsoló fehérje második huroknál (LPL2) ember-sertés-patkány-csirke, addig a harmadik huroknál (LPL3) ember-patkány-sertés-csirke a sorrend a dendrogram alapján. Feltűnő még, hogy a versican második hurka (HVERSL2) a várttal ellentétben nagyobb hasonlóságot mutat a G1 globuláris egység második hurkához (PGG1L2), mint a G2 globuláris egység második hurkához (PGG2L2). Ez is génkonverziós jelenségre utal, mivel mind a versican első hurok (HVERSL1), mind a versican harmadik hurok (HVERSL3) esetében nyilvánvaló, hogy az elválás az előzőekben ismertetett modellünk szerint történt.

ÖSSZEFOGLALÁS

A kapcsoló fehérje a porcszövetben stabilizálja a proteoglükán hialuronsav aggregátumokat azáltal, hogy kötődik mindkét makromolekulához.

Maga a kapcsoló fehérje három modulból áll, melyeket külön exonok kódolnak. Az exonokat egyes fázisú intronok szakítják meg. Ez jellemző a többi mozaik fehérjére is, melyekben a kapcsoló fehérjében találhatókhoz homolog modulok vannak. Egy negyedik exon kódolja a prekurzor fehérje szekréciós szignál peptidjét. A korábban azonosított genomi klónok azonban nem tartalmazták a gén 5' végét. Ezért tűztük ki célul a hiányzó első exont hordozó cDNS klónok kinyerését és segítségükkel a gén 5' végének az azonosítását. A rokon moduloknak az összehasonlító vizsgálata pedig segít a mozaik fehérjék kialakulásának a megértésében. Munkánk során a következő eredményeket értük el:

1. Specifikus oligonukleotid primerek segítségével sikerült egy, a kapcsoló fehérje gén 5' végére jellemző cDNS könyvtárat létrehoznunk. A pozitív cDNS klónokat jellemeztük és meghatároztuk a szekvenciájukat. Igazoltuk, hogy a leghosszabb cDNS klón 5' vége is kapcsoló fehérje gén eredetű.
2. Izoláltunk egy genomiális klónt, mely tartalmazza a gén promoter régióját, az első exont és egy nagy darabot az első intronból. Meghatároztuk ennek a klónnak a restrikciós térképét, valamint a promoter régió és az első exon szekvenciáját. Bebizonyítottuk, hogy a klón nem fed át a már régebben izolált, a második exont hordozó genomiális klónokkal, így annyit tudunk mondani, hogy a kapcsoló fehérje gén mérete meghaladja a 100 kb-t.
3. Kétféle módszerrel, primer extenzióval és T4 polimerázos térképezéssel meghatároztuk a transzkripció kezdőpontjait. Megállapítottuk, hogy a mRNS átirása legalább öt pontból indul.
4. A cDNS klónok és a genomiális klónok összehasonlításával megállapítottuk, hogy az mRNS 5' nem transzlálódó részén alternatív exonhasználat történik. E jelenséget PCR kisérletekkel igazoltuk. Az izolált cDNS klónok alapján legalább ötféle, az 5' végében különböző mRNS típus létezik, ezenkívül a PCR eredményeképp még legalább háromféle típus létezését feltételezzük.
5. A különböző cDNS formák szekvenciájának az analízise feltárta, hogy egyes mRNS-ek másképp viselkedhetnek a transzláció iniciációjának szempontjából. A 7G12 és a 11B4 típusú mRNS-ekben található az iniciációs AUG kódontól 5' irányban két nyitott leolvasási keret, ami erősen gátolhatja a transzláció megkezdését. A fehérjeszintézis

szempontjából fontos, hogy milyen másodlagos szerkezetet vesz fel az mRNS 5' vége. A különböző mRNS formák ebből a szempontból is eltérőnek mutatkoznak a másodlagos szerkezet predikcióink alapján.

6. Primer extenzióval, irányított cDNS klónozással PCR kisérletekkel próbáltuk megnézni, hogy létezik-e a csirkében is a patkány kondroszarkóma kapcsoló fehérjében leirt alternatív exon a második és a harmadik exon között. Valamennyi ezirányú kisérletünk negatív eredménnyel zárult, tehát valószínüleg a csirkében hiányzik ez az alternatív exon.

A gén 5' végének, a promoter régiónak a klónozása és további vizsgálata lehetőséget teremt ahhoz, hogy megértsük a kapcsoló fehérje expressziójának a szabályozását. Ez azért is érdekes, mivel alkalmat ad a porcszövetben termelődő más extracelluláris fehérjék (II. típusú kollagén, CMP) génjeinek szabályozó ellemeivel történő összehasonlításra.

7. Összegyűjtöttük egyrészt a kapcsoló fehérje első hurokkal rokon modulokat, másrészt a második és a harmadik hurokkal rokon modulokat és összehasonlítottuk őket. A rokon modulok rokonsági kapcsolatait dendrogrammal ábrázoltuk. A dendrogramok alapján egy modellt állítottunk fel, mellyel magyarázatot próbálunk adni az egy fajban található tízféle PTR és háromféle első hurok modul kialakulására egy ős PTR és ős első hurok modulból. Hipotézisünk szerint a mai napig megismert PTR domének kialakulása hat független duplikáció során történt, ebből három duplikációt egy első hurok és két PTR modul egyszerre szenvedett el.

A dolgozat anyagából készült publikációk

1. Deák, F., Barta, E., Mestric, S., Biesold, M. és Kiss, I. (1991) Complex pattern of alternative splicing generates unusal diversity in the leader sequence of the chicken link protein mRNA. *Nucleic Acids Res.* **19**, 4983-4990.

2. Barta, E., Deák, F., Kiss, I. (1993) Evolution of the hyaluronan-binding module of link protein. *Biochem. J.* megjelenés alatt.

Egyéb publikáció

1. Kiss, I., Bősze, Zs., Szabó, P., Altanchimeg, R., Barta, E. és Deák, F. (1990) Identification of positive and negative regulatory regions controlling expression of the cartilage matrix protein gene. *Mol. Cell. Biol.* **10**, 2432-2436.

IRODALOM

1. Heinegård, D., and Paulsson, M. (1987) *Methods Enzymol.* **145**, 336-363.
2. Caplan, A.I. (1984) *Sci. Am.* **10/1984**, 82-90.
3. *Cartilage*, Vols. 1, 2, 3, szerk. Hall, B.K., (1983) Academic Press.
4. Stockwell, R.R., *Biology of cartilage cells*, (1979) Cambridge University Press.
5. *Cartilage: Molecular aspects*, szerk. Hall, B., Newman, S. (1991) CRC Press.
6. Sajdera, S.W., Hascall, V.C. (1969) *J. Biol. Chem.* **244**, 77-87.
7. Hascall, V.C., Sajdera, S.W. (1969) *J. Biol. Chem.* **244**, 2384-2396.
8. Heinegard, D. (1972) *Biochim. Biophys. Acta* **285**, 181-192.
9. Smith, B.D., Martin, G.R., Miller, E.J., Dorfman, A. és Swarm, R. (1975) *Arch. Biochem. Biophys.* **166**, 181-186
10. Eyre, D.R. (1985) *Science* **207**, 1315-1322.
11. Bornstein, P. és Sage, H. (1980) *Ann. Rev. Biochem.* **49**, 957-1003.
12. Miller, E.J. és Matukas, V.J. (1969) *Proc. Natl. Acad. Sci. USA* **64**, 1264-1268.
13. Reese, C.A. és Mayne, R. (1981) *Biochemistry* **20**, 5443-5448.
14. van der Rest, M., Rosenberg, L.C., Olsen, B.R. és Poole, A.R. (1986) *Biochem. J.* **237**, 923-925.
15. Müller-Glauser, W., Humbel, B., Glatt, M., Strauli, P., Winterhalter, K.H. és Bruckner, P. (1986) *J. Cell Biol.* **102**, 1931-1939.
16. Heinegård, D., Oldberg, A. (1989) *FASEB J.* **3**, 2042-2051.
17. Heinegard, D., Sommarin, Y. (1987) *Methods Enzymol.* **144**, 305-319.
18. Doege, K., Sasaki, M., Horigan, E., Hassell, J.R., Yamada, Y. (1988) *J. Biol. Chem.* **262**, 17757-17767.
19. Halberg, D.F., Proulx, G., Doege, K., Yamada, Y. és Drickamer, K. (1986) *J. Biol. Chem.* **261**, 8108-8111.
20. Hascall, V.C., Hascall, G.K. (1981) *in Cell biology of extracellular matrix*, szerk. Hay, E.D., Plenum, N.Y. pp.39-63.
21. Hascall, V.C., Sajdera, S.W. (1970) *J. Biol. Chem.* **245**, 4920-4930.
22. Hardingham, T., Muir, H. (1972) *Biochim. Biophys. Acta* **279**, 401-405.
23. Rouslahti, E. (1988) *Annu. Rev. Cell. Biol.* **4**, 229-255.
24. Fischer, L., Termine, J. és Young, M. (1989) *J. Biol. Chem.* **264**, 4571-4576.

25. Tsonis, P.A. és Goetinck, P.F. (1988) *Exp. Eye Res.* **46**, 753-764.

26. Argraves, W.S., Deák, F., Sparks, K.J., Kiss, I. és Goetinck, P.F. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 464-468.

27. Kiss, I., Deák, F., Holloway, R.G., Jr., Delius, H., Mebust, K.A., Frimberger, E., Argraves, W.S., Tsonis, P.A., Winterbottom, N., Goetinck, P.F. (1989) *J. Biol. Chem.* **264**, 8126-8134.

28. Kiss, I., Bősze, Zs., Szabó, P., Altanchimeg, R., Barta, E. és Deák, F. (1990) *Mol. Cell. Biol.* **10**, 2432-2436.

29. Hedbom, E. és Heinegård, D. (1989) *J. Biol. Chem.* **264**, 6898-6905.

30. Keiser, H., Schulman, H.J. és Sandson. J.I. (1972) *Biochem. J.* **126**, 163-169.

31. Baker, J.R. és Caterson, B. (1979) *J. Biol. Chem.* **254**, 2387-2395.

32. Roughley, P.J., Poole, A.R. és Mort, J.S. (1982) *J. Biol. Chem.* **257**, 11908-11915

33. LeGlédic, S., Périn, J.P., Bonnet, F. és Jollès, P. (1983) *J. Biol. Chem.* **258**, 14759-14765.

34. Gregory, J.D. (1973) *Biochem. J.* **133**, 383-386.

35. Heinegard, D., és Hascall, V.C. (1974) *J. Biol. Chem.* **249**, 4250-4256.

36. Hardingham, T.E. (1979) *Biochem. J.* **177**, 237-247.

37. Heinegard, D., és Hascall, V.C. (1979) *J. Biol. Chem.* **254**, 921-926.

38. Franzen, A., Björnsson, S. és Heinegard, D. (1981) *Biochem. J.* **197**, 669-674.

39. Heinegard, D. és Hascall, V.C. (1979) *J. Biol. Chem.* **254**, 927-934.

40. Tengblad, A. (1981) *Biochem. J.* **199**, 297-305.

41. Baker, J.R. és Neame, P.J. (1987) *Methods Enzymol.* **144**, 401-412.

42. Gardell, S., Baker, J.R., Caterson, B., Heinegård, D. és Rodén, L. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1832-1839.

43. Vijayagopal, P., Radhakrishnamurthy, B., Srinivasan, S.R. és Berenson, G.S. (1985) *Biochim. Biophys. Acta* **839**, 110-121.

44. Poole, A.R., Pidoux, I., Reiner, A., Cöster, L. és Hassel, J.R. (1982) *J. Cell Biol.* **93**, 910-917.

45. Fife, R.S., Caterson, B. és Myers, S.L.(1985) *J. Cell Biol.* **100**, 1050-1058.

46. Deák, F., Kiss I., Sparks K.J., Argraves W.S., Hampikian, G. és Goetinck P.S. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3766-3770.

47. Périn, J.P., Bonnet, F., Pizon, V., Jollès, J. és Jollès, P. (1980) *FEBS Let.*. **119**, 333-343.

48. Baker, J.R., Caterson, B. és Christner, J.E. *Limb Development and Regulation.* (1983) szerk. Kelley, R.O., Goetinck, P.F. és MacCabe, J.A. Liss Press.

49. Treadwell, B.J., Mankin, D.P., Ho, P.K. és Mankin, H.J. (1980) *Biochemistry* **19**, 2269-2273.

50. Neame, P.J., Christner, J.E. és Baker, J.R. (1986) *J. Biol. Chem.* **261**, 3519-3535.

51. Doege, K., Hassel, J.R., Caterson, B. és Yamada, Y. (1986) *J. Biol. Chem.* **261**, 3519-3535

52. Perkins, S.J., Nealis, A.S., Dudhia, J. és Hardingham, T.E. (1989) *J. Mol. Biol.* **206**, 737-753.

53. Dudhia, J. & Hardingham, T.E. (1990) *Nucleic Acids Res.* **18**, 1292.

54. Kiss, I., Deák, F., Mestric, S., Delius, H., Soós, J., Dékány, K., Argraves, W.S., Sparks, K.J. és Goetinck P.F. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 6399-6403.

55. Rhodes, C.,K. Doege, K., M. Sasaki, M. és Yamada, Y.(1988) *J. Biol. Chem.* **263**, 6063-6067.

56. Gilbert, W. (1978) *Nature* **271**, 501.

57. Traut, T.W. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2944-2948.

58. Dorit, R.L., Schoenbach, L. és Gilbert, W. (1990) *Science* **250**, 1377-1381.

59. Patthy, L. (1987) *FEBS Let.* **214**, 1-7.

60. Patthy, L. (1985) *Cell* **41**, 657-663.

61. Bork, P. (1991) *FEBS Let.* **286**, 47-54.

62. Périn, J.-P., Bonnet, F., Thurieau, C. és Jolles, P. (1987) *J. Biol. Chem.* **262**, 13269-13272.

63. Bonnet, F., Perin,, J.-P., Lorenzo, F., Jolles, J. és Jolles, P. (1986) *Biochim. Biophys. Acta* **873**, 152-155.

64. Doege, K.J., Sasaki, M., Kimura, T. és Yamada, Y. (1991) *J. Biol. Chem.* **266**, 894-902.

65. Chandrasekaran, L. és Tanzer, M.L. (1992) *Biochem. J.* **288**, 903-910

66. Zimmermann. D. és Ruoslahti, E. (1989) *EMBO J.* **8**, 2975-2981.

67. Rauch, U., Karthikeyan, L., Maurel, P., Margolis, R.U. és Margolis, R.K. (1992) *J. Biol. Chem.* **267**, 19536-19547.

68. Springer, T.A. (1990) *Nature* **346**, 425-434.

69. Wiedemann, H., Paulsson, M., Timkpl, R., Engel, J. és Heinegård, D. (1984) *Biochem. J.* **224**, 331-333.

70. Paulsson, M., Morgelin, M., Wiedemann, H., Beardmore-Gray, M., Dunham, D., Hardingham, T., Heinegård, D., Timpl, R. és Engel, J. (1987) *Biochem. J.* **245**, 763-772.

71. Perides, G., Lane, W.S., Andrews, D., Dahl, D. és Bignami, A. (1989) *J. Biol. Chem.* **264**, 5981-5987.

72. Bignami, A., Lane, W.S., Andrews, D. és Dahl, D. (1989) *Brain Res. Bul.* **22**, 67-70.

73. Perides, G., Biviano, F. és Bignami, A. (1991) *Biochim. Biophys. Acta* **1075**, 248-258.

74. Perris, R. és Johansson, S. (1990) *Dev. Biol.* **137**, 1-12.

75. Marks, M.S., Chi-Rosso, G. és Toole, B.P. (1990) *J. Neurochem.* **54**, 171-180.

76. Goldstein, L.A., Zhou, D.F.H:, Picker, L.J., Minty , C.N., Bargatze, R.F., Ding, J.F. és Butcher, E.C. (1989) *Cell* **56**, 1063-1072.

77. Stamenkovic, I., Amiot, M., Pesando, J.M. & Seed, B. (1989) *Cell* **56**, 1057-1062.

78. Lee, T.H., Wisniewski, H.-G. és Vilcek, J. (1992) *J. Cell Biol.* **116**, 545-557.

79. Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C.B. és Seed, B. (1990) *Cell* **61**, 1303-1313.

80. Günthert, U., Hofmann, M., Rudy, W., Reber, S., Zöller, M., Hauβmann, I., Matzku, S., Wenzel, A., Ponta, H. és Herrlich, P. (1991) *Cell* **65**, 13-24.

81. Sinha, N.D., Biernat, J., McManus, J. és Köster, H. (1984) *Nucleic Acids Res.* **12**, 4539-4557.

82. Sambrook, J., Fritsch, E.F. és Maniatis, T. (1989) *Molecular cloning, a laboratory manual.* 2nd ed., CSH Laboratory Press.

83. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557-580.

84. Vieira, J. és Messing, J. (1982) *Gene* **19**, 259-268.

85. Messing, J. (1983) *Methods Enzymol.* **101**, 20-78

86. Holmes, D.S. és Quigley, M. (1981), *Anal. Biochem.* **114**, 193

87. Clewell, D.B. és Helinski, D.R. (1972) *J. Bacteriol.* **110**, 1135-1141.

88. Tanaka, T. és Weisblum, B.J. (1975) *J. Bacteriol.* **121**, 354-362.

89. Feinberg, A. P. és Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13.

90. Maxam, A., Gilbert, W. (1980) *Methods Enzymol.* **65**, 499-560.

91. Adams, S.L., Sobel, M.B., Howard, B.H., Olden, K., Yamada, K.M., de Crombrugghe, B. és Pastan, I. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3399-3403.

92. Aviv, H. és Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408-1412.

93. Hanahan, D. és Meselson, M. (1983) *Methods Enzymol.* **100**, 333-342.

94. Sanger, F., Nicklen, S. és Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467

95. Hu, M. C-T. és Davidson, N. (1986) *Gene* **42**, 21-29

96. Devereux, J., Haeberli, P. és Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387-395.

97. Zuker, M. és Stiegler, P. (1981) *Nucleic Acids Res.* **9**, 133-148.

98. Shapiro, M.B. és Senapathy (1987) *Nucleic Acids Res.* **15**, 7155-7175.

99. Bucher, P. és Trifonov, E.N. (1986) *Nucleic Acids Res.* **14**, 10009-10026.

100. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125-8148.

101. Antonsson, P., Heinegård, D. és Oldberg, Å. (1989) *J. Biol. Chem.* **264**, 16170-16173.

102. Nottenburg, C., Rees, G. és John, T.S. *Proc. Natl. Acad. Sci. USA* **86**, 8521-8525.

103. Idzerda, R.L., Carter, W.G., Nottenburg, C., Wayner, E.A., Gallatin, W.M. és John, T.S. *Proc. Natl. Acad. Sci. USA* **86**, 4659-4663.

104. Bosworth, B.T., Dowbenko, D., Schuster, D.E., Jutila, M. és Harp, J.A. (1991) *Mol. Immunol.* **28**, 1131-1135

105. Goetinck, P.F., Stripe, N.S., Tsonis, P.A. és Carlone, D. *J. Cell Biol.* (1987) **105**, 2403-2408.

106. Rhodes, C., Savanger, P., Line, S., Sasaki, M., Chirigos, M., Doege, K. és Yamada, Y. (1991) *Nucleic Acids Res.* **19**, 1933-1939.

107. Kozak, M. (1988) *J. Cell Biol.* **107**, 1-7.

108. Lowe, W.L., Roberts, C.T., Lasky, S.R. és LeRoith, D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 8946-8950.

109. López-Casillas, F., Luo, X., Kong, I-S. és Kim, K-H. (1989) *Gene* **83**, 311-319.

110. López-Casillas. F. és Kim, K-H. (1989) *J. Biol. Chem.* **264**, 7176-7184.

111. Peralta, E.G., Winslow, J.W., Peterson, G.L., Smith, D.H., Ashkenazi, A., Ramachandran, J., Schimerlik, M.I. és Capon, D.J. (1987) *Science* **236**, 600-605.

112. Deák, F., Barta, E., Mestric, S., Biesold, M. és Kiss, I. (1991) *Nucleic Acids Res.* **19**, 4983-4990.

113. Kozak, M. (1986) *Cell* **44**, 283-292.

114. Kozak, M. (1987) *J. Mol. Biol.* **196**, 947-950.

115. Kozak, M. (1989) *Mol. Cell Biol.* **9**, 5073-5080.

116. Kozak, M. (1989) *J. Cell Biol.* **108**, 229-241

117. Hawkins, J.D. (1988) *Nucleic Acids Res.* **16**, 9893-9908.

118. Rouault, T.A., Hentze, M.W., Gaughman, S.W., Harford, J.B. és Klausner, R.D. (1988) *Science* **241**, 1207-1210.

119. Kozak, M. (1989) *Mol. Cell Biol.* **9**, 5134-5142.

120. Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2850-2854.

121. Kozak, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 8301-8305.