

Élő Péter

Egy atadenovírus izolátum genomjának
összehasonlító elemzése

Ph.D. értekezés

Témavezetők: dr. Benkő Mária és dr. Harrach Balázs

MTA Állatorvostudományi Kutatóintézete

Budapest
2002

A dolgozatban használt rövidítések jegyzéke

ADP	– adenovírus halál (death) protein
ATF	– adenovírus gazdafaktor
as	– aminosav
II	– hexon fehérje
BAdV	– szarvasmarha adenovírus
bp	– bázispár
CAdV	– kutya adenovírus
CAR	– coxsackie vírus és adenovírus receptor
CELO	– FAdV–1
CPE	– citopatikus hatás
CTF1	– transzkripció faktor
DAdV–1	– kacsa 1-es adenovírus
DBP	– DNS-kötő fehérje
EDS	– kacsa 1-es adenovírus
EGF-R	– epidermális növekedési faktor
eIF2	– eukarióta iniciációs faktor 2
FAdV–	– madár adenovírus
FrAdV–	– béka adenovírus
HAdV	– emberi adenovírus
ICTV	– International Committee on Taxonomy of Viruses
ITR	– fordított végi ismétlődés
kbp	– kilobázis-pár
L	– késői gének mRNS-einek jelzője
LH	– adenovírus genom bal oldali régiója
MAdV	– egér adenovírus
MEM	– módosított Eagle féle tápoldat
MHC–	– fő hisztokompatibilitási faktor
MLP	– késői promóter
NCBI	– National Center for Biotechnology Information
NFI	– sejtmagi (transzkripció) faktor I
NFIII	– sejtmagi (transzkripció) faktorIII
NLS	– mágis lokalizációs szignál
OAV287–	– juh adenovírus 287-es izolátum
ORF	– nyílt leolvasási keret
PCR	– polimeráz láncreakció
pol	– DNS-polimeráz
SASP	– kisméretű, savoldékony protein (Gram-pozitív baktériumok)
TAdV–	– pulyka adenovírus

TNF – tumor nekrosis faktor
TP – terminális protein
TPL – három részből álló szekvencia a kései gének ORF-jei előtt
VA RNS– vírus-kapcsolt RNS

Tartalomjegyzék

1. Irodalmi áttekintés	8
1.1. Az adenovírus kutatás története	8
1.2. Morfológia	10
1.3. Vírusszaporodás	11
1.4. Taxonómia	13
1.4.1. Az adenovírusok javasolt rendszere	15
1.5. Genomszerveződés	16
1.5.1. Mastadenovírusok	16
1.5.2. Atadenovírusok	16
1.5.3. Aviadenovírusok	17
1.5.4. Siadenovírusok	17
1.6. Célkitűzések	18
2. Anyagok és módszerek	20
2.1. Felhasznált vegyszerek és enzimek	20
2.2. Molekuláris és mikrobiológiai technikák	20
2.2.1. Vírusszaporítás	20
2.2.2. A vírus-DNS tisztítása	20
2.2.3. PCR reakciók	21
2.2.4. Molekuláris klónozás	21
2.2.5. DNS-szekvenálás	23
2.3. A DNS-szekvenciák szerkesztése, analízise és ábrázolása	23
2.3.1. Filogenetikai számítások	24
3. Eredmények és megbeszélés	26
3.1. A Rus törzs genomszerveződése	26

3.2. Az ITR szekvenciák	27
3.3. Korai régiók	31
3.3.1. A bal oldali korai régió	31
3.3.2. Az „E3” régió fehérjéi	33
3.3.3. Az E4 régió fehérjéi	34
3.3.4. A DNS-replikáció vírus eredetű faktorai	41
3.3.5. A késői fehérjék	42
3.3.6. Struktúrális fehérjék	45
4. Összefoglalás	56
5. Summary	58
6. Köszönetnyilvánítás	61
A. Az illesztésekhez használt fehérjék génbanki elérhetősége	75

Ábrák jegyzéke

1.1. Mastadenovírus vázlatos rajza	10
1.2. HAdV-40 géntérképe	16
1.3. OAV287 géntérképe	17
1.4. FAdV-1 géntérképe	17
1.5. FrAdV-1 géntérképe	18
3.1. A Rus adenovírus vázlatos géntérképe	26
3.2. ITR szekvenciák	29
3.3. Az LH2 proteinek illesztése	32
3.4. Az LH3 proteinek illesztése	36
3.5. A IVa2 mRNS-ének kódoló szakaszán feltételezett splicing. . .	37
3.6. Atadenovírusok E3 fehérjéinek illesztése	37
3.7. A genom jobb végének duplikációi	38
3.8. Az „E3” szakaszon felismerhető homológ DNS szakaszok	38
3.9. Az „E3” szakasz fehérjéinek „mag” régiójából számított törzsfa	39
3.10. Az E4 34K fehérjék illesztése	40
3.11. A pTP mRNS splicingja	41
3.12. A DNS-kötő fehérjék C-terminálisának illesztése	42
3.13. A 33K protein mRNS-ének kódoló szakaszán feltételezett splicing	43
3.14. A proteáz enzimek törzsfája	44
3.15. p32K fehérjék illesztése	45
3.16. p32K-SASP illesztés	46
3.17. Jobb p32K-SASP illesztés	47
3.18. Felerősített p32K génrészletek	49
3.19. p32K törzsfa	50
3.20. A hexon fehérjékből készített törzsfa	51
3.21. Fiber fehérjék farki részének illesztése	52

3.22. A Rus fiber felépítése	53
3.23. Néhány pVII protein illesztése	55

Táblázatok jegyzéke

3.1. A Rus vírus fehérjéi	28
-------------------------------------	----

1. fejezet

Irodalmi áttekintés

1.1. Az adenovírus kutatás története

Az adenovírusokat Rowe és munkatársai fedezték fel 1953-ban,¹ mikor gyermekek műtéti úton eltávolított manduláiból próbáltak szövettényészetet készíteni, de azt tapasztalták, hogy valamiféle átvihető ágens az epitheliális sejtek degenerációját okozza. Hamarosan további humán adenovírusokat izoláltak különböző légzőszervi bántalomban szenvedő betegekből, és már röviddel az első izolálások után számos szerotípust lehetett elkülöníteni.

Bár az izolált humán adenovírusoknak csak mintegy harmada okoz megbetegedést, több, korábban már leírt betegségről is kiderült, hogy adenovírus-fertőzéssel hozható kapcsolatba. Ezek közül legismertebbek a járványos köthártyagyulladás² és a gyermekkori virális eredetű hasmenéses megbetegedések jelentős része.³

Az adenovírusok molekuláris biológiájának kutatása akkor került előtérbe, amikor 1962-ben felfedezték, hogy bizonyos humán adenovírusok rácsálókba oltva tumorokat indukálnak.⁴ A felfedezést követő években az intenzív kutatásoknak köszönhetően az adenovírusok „modellszervezetté” váltak a DNS replikáció, a transzkripció, az mRNS érési folyamatok, a polipeptidok sejten belüli transzportjának és a nukleoprotein komplexek képződésének vizsgálatában. Számos azóta is használt molekuláris biológiai kísérleti eljárást az adenovírus gének transzkripciójának vizsgálatára dolgoztak ki. E vizsgálatok során fedezték fel az RNS-splicing mechanizmusát is.⁵

Miután a leggyakoribb humán adenovírusok (HAdV) génjei és azok működése ismertté váltak, fölmerült vektorként való felhasználásuk igénye vakcinázási, fehérje-kifejezési és génterápiás kísérletekben^{6;7} is. Noha a kezdeti túlzó remények – főként az immunrendszer adenovírusokkal szemben meg-

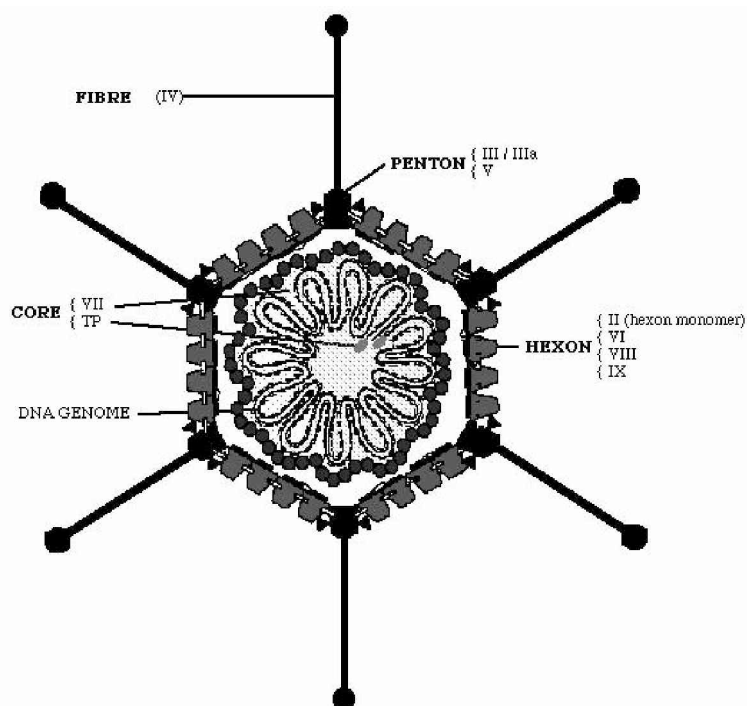
nyilvánuló igen hatékony válasza miatt⁸ – nem váltak valóra, napjainkban már kereskedelmi forgalomban is kapható rekombináns fehérjék előállítására alkalmas, HAdV-5-ön alapuló vektorrendszer (pl.: Qbiogen: AdenoVator).

A vizsgálatokat azonban csupán néhány jól szaporítható humán vírusról végezték, míg az állati eredetű adenovírusok kutatása lényegesen lassabban haladt. Meglepő módon az első madár adenovírust már 1949-ben megtalálták, de tévesen egy szarvasmarha bőrbetegség okozójának vélték.⁹ Azonosítása 1957-ben történt „chicken embryo lethal orphan” (CELO) néven.¹⁰ 1965 után, mikor kiderült róla, hogy némely humán eredetű adenovírushoz hasonlóan, újszülött rágcsálókba oltva tumorokat indukálhat, egy ideig intenzív vizsgálatok tárgya volt. Az első szarvasmarha adenovírust 1959-ben írták le egy olyan vizsgálat keretében, ami a humán vírusok szarvasmarhákban előforduló megfelelőit célozta azonosítani.¹¹ Az 1960-as évek közepére beteg és látszólag egészséges szarvasmarhákból is izoláltak olyan adenovírus törzseket, amelyeket szerológiai tulajdonságaik, hővel és gyenge savakkal szembeni viszonylag ellenálló voltak és rossz szaporíthatóságuk miatt Bartha új alcsoportba sorolt¹². További vizsgálatok alapján javaslatot tettek új genusba sorolásukra is,^{13;14} de ezt az ICTV^a „Adenovírus Munkacsoportja” akkor elutasította. A tojó tyúktojástermelésének járványos csökkenését vizsgálva fedezték fel az „egg drop syndrome” (EDS) vírust, melynek a hivatalos neve ma az egyes típusú kacsák adenovírus, rövidítése pedig DAdV-1. A DNS szekvenálás elterjedéséig az állati adenovírusoknál az esetleírásokon túl többnyire megelégedtek a vírusok szerológiai jellemzésével, vagy új gazdafajnál elektronmikroszkópos azonosítással. Emiatt, noha akkoriban már az összes gerinces osztályból sikerült adenovírusokat azonosítani, – beleértve a halakat¹⁵ és a hüllőket is¹⁶ – megbízható vírusjellemezés híján a szisztematikai felosztás sokáig nem tudta követni a tényleges állapotot. Sőt a hivatalosan elfogadott rendszertani felosztás a mai napig nem tükrözi a hetvenes évek óta jócskán megváltozott ismeretanyagot.

Ahogy a tömeges DNS-szekvenálás a nyolcvanas évek vége felé elterjedt, egyre több szekvenciárészlet vált ismertté, illetve egy sor vírus teljes genomsekvenciáját is meghatározták. Amíg az első humán adenovírus genomsekvencióját (HAdV-2¹⁷) 1986-ban, az első „atipikus mastadenovírusét” (OAV287¹⁸) és az első aviadenovírusét” (FAdV-1¹⁹) csak 1996-ban, az első „atipikus aviadenovírusét” (TAdV-3²⁰) 1998-ban fejezték be, az első tipikus állati mastadenovírusét pedig (CAdV-1²¹) 1997-ben. Jelenleg 21 adenovírus teljes DNS-szekvenciája ismert.

^aNemzetközi Virustaxonómiai Bizottság: International Committee on the Taxonomy of Viruses

1.2. Morfológia



1.1. ábra. Mastadenovírus vázlatos rajza^b

^bA képet a „<http://www.tulane.edu/~dmsander/WWW/335/Adenoviruses.html>” internet-oldalról vettem át.

Az adenovírusok virionjai 70–90 nm átmérőjű, burok nélküli, ikozaéderes részecskék, mintegy 23% DNS-t és 77% fehérjét tartalmaznak. A kapszid csúcsairól egy vagy két antenna-szerű, fibernek nevezett nyúlvány ered, melynek hossza szerotípustól függően 9–77,5 nm, és a disztális végén gomb-szerű, fiber fejnek is nevezett képletben végződik (1.1 ábra). A szárát és gombot három speciális tripla β -spirálba rendeződött polipeptid alkotja.²² A fiberek szárán a fajra, míg a fejen az adott típusra jellemző antigén epitopok vannak. A kapszid 240 lap kapszomerből (hexon) és 12 csúcsi kapszomerből (penton) áll. Az utóbbi fehérjének önmagában is toxikus hatása van a sejtekre.²³ A csúcsi kapszomereken az adott alcsoportra jellemző epitopokat azonosítottak. A hexonok mindegyikét három azonos, II-es proteinnek nevezett polipeptid építi fel. Alapi részükön nemzetségspecifikus, a külső oldalukon pedig típusspecifikus epitopok találhatók. A pentonok a fibert alkotó három polipeptidből

(IV), és a hozzájuk szorosan kapcsolódó penton-alapból állnak, melyet öt III-as protein épít fel. A kapszid éleinél 2–2 IIIa fehérje monomer helyezkedik el. A VI-os protein a pentonok körüli hexonok alatt képez több polipeptidből álló gyűrűket. A VIII, a VI, a IIIa és a IX fehérjék a hexonok alatt, azokkal együtt egységes réteget alakítanak ki. Az V, VII, X számú fehérjék, valamint a terminális fehérje (TP) a DNS-hez kötődve alkotják a magszerkezetet.⁸

1.3. Vírusszaporodás

Az alábbi részben a mastadenovírusokra általánosan jellemzőnek tartott mechanizmusokat tekintem át röviden, mivel a humán adenovírusokon végzett molekuláris biológiai vizsgálatok folytán erről a csoportról pontos ismereteink vannak. Az adenovírusok általában szűk gazdaspektrummal rendelkező vírusok, az eredeti gazdáról más fajra való átvitelük ritkán következik be, amint azt a BAdV-2^{24;25} esetében leírták. Az adenovírus a fiber „gomb”-szerű részével tapad meg a sejtreceptoron. Ez a részletesen vizsgált emberi adenovírusok esetében a „coxsackie vírus és adenovírus receptor” (CAR). (Ennek a folyamatnak szerepe van a virális proteáz aktiválásában.²⁶) Ezután a penton-alap és a sejt integrinjeinek kapcsolódása elősegíti az endoszomákba kerülést, és a vírus a normális endocitotikus utat követve bejut a sejtbe. A fiberek ekkor válnak le.²⁷ A kapszid ezután fokozatosan fellazul, és később, már a citoszolban, szétesik.²⁸ Ebben a folyamatban fontos szerepe van az endoszomás pH csökkenésének, ami a virális proteáz aktiválódásának másik kiváltó oka. A kapszidban lévő proteáz hasítja a hexonokat a vírusmaghoz (core) kapcsoló VI-os proteint, az endoszomás proteázok pedig részlegesen hasítják a kapszid külső fehérjéit. Ez a folyamat lazítja fel a kapszidot, ami kijut a citoplazmába, ahol leválnak a IIIa, VIII és IX fehérjék. Amikor a vírus eléri a sejtmag-membránt, a vírusmag, melyet a DNS, a VII, a μ és a lazábban kapcsolódó V-ös fehérje alkot, eddig ismeretlen módon kikerül a citoplazmába. Innen valószínűleg az V fehérje segítségével a citoszkeletonhoz kapcsolódó transzportrendszerhez kötődve jut be a magba.²⁹ Az itt zajló struktúrális változásokról keveset tudunk, de annyi bizonyos, hogy a vírust alkotó fehérjék pl. a DNS-kötő fehérje (DBP^c), IVa2, V eloszlása jellegzetes változásokon megy keresztül.^{30;31;32}

A HAdV-2 vírus transzkripciója kilenc promoter felhasználásával zajlik, ezek közül egyetlen promoterről átíródó mRNS nem szerkesztett, a pIX-é. Az összes többi több exonból épül fel, és a IVa2 kivételével mindegyikről többféle mRNS képződik.

^cDNA binding protein

A korai gének kifejeződése az E1A és E4 fehérjék expressziójával kezdődik. Ezeket „immediate early” géneknek is hívják, szerepük többek között a többi korai gén aktiválása, és bizonyos sejtleletani folyamatok szabályozása.⁸ A korai gének közé tartoznak a főként az immunrendszert moduláló fehérjéket kódoló gének az E3 régióban, és a DNS replikálásáért felelős három vírusfehérje: a DNS-polimeráz (pol), a DBP, és a terminális protein (TP) is, melyek az E2 régió termékei. A korai gének átírása több promoterről folyik, ezek száma és elhelyezkedése nemzetségenként különböző. Az E3 gének által kódolt fehérjék közé igen különböző hatású fehérjék tartoznak. Az átírt fehérjék funkció szempontjából is csak részben ismertek.³³ A Gp19K a leggyakoribb korai fehérje,³⁴ az endoplazmatikus retikulum membránjába ékelődve az MHC-1 fehérjéhez kötődik,³⁵ és akadályozza azok transzportját a sejtmembránba.³⁶ A 14.7K protein a TNF^d-on keresztül akadályozza a sejtízist.³⁷ A 10.4K protein az EGF-R^e bontásának gyorsításával serkenti a sejtanyagcserét.³³ Az E3-11.6K protein, melyet ADP^f-nek is neveznek,³⁸ egy transzmembrán glikoprotein, mely a sejt lízisében játszik fontos szerepet.

Az E4 fehérjék közül az ORF3-ról és ORF4-ről kifejeződő fehérjék valamelyikének megléte^{39;40} nélkülözhetetlen a vírusok sejttényészen való szaporíthatóságához.

A humán adenovírusokra jellemző, általuk termelt VA RNS egy kinázhoz kötődik, ezzel megakadályozza a transláció indításában kulcsszerepet játszó eIF2 foszforilálását, ami így nem képes a fehérjeszintézis leállítására.⁴¹ Később pedig a riboszomákhoz kötődve segíti a virális fehérjék translációját.⁴²

A DNS replikáció megindulásától számítjuk a késői szakaszt, melyben a vírus struktúr-fehérjéinek szintézise és a vírusrészecskék összeépülése történik. Ekkorra az E1B 55K protein az E4 régió fehérjéivel leállítja a sejteredetű mRNS-ek transzportját a citoplazmába.³⁹ Az E1B 55K–E4 Orf6 komplex megkezdí a vírus eredetű mRNS-ek citoplazmába szállítását,⁴³ közben az mRNS-ek splicingjának mintázata – részben szintén az E4 proteinek hatására is – változásokon megy keresztül. A replikációhoz a vírusfehérjéken kívül még két gazda-faktor is szükséges: az egyik egy topoizomeráz,⁴⁴ a másik egy transzkripciós faktor (CTF1).⁴⁵ A DNS szintézise mindkét szálon folyamatosan zajlik, így tehát az egyik szál mindaddig replikációsan inaktív és DBP-vel borított, amíg a másik szál mellett folyó szintézis teljesen fel nem szabadítja⁴⁶ azt.

A kései gének mindegyike egyetlen promoterről, a késői fő promoterről

^dtumor necrosis factor

^eepidermal growth factor

^fAdenovirus death protein

(MLP)^g íródik át. A transzkriptumok terminálása azonban különböző helyekenél történik. Az érett mRNS-ek alternatív splicing folyamatok során jönnek létre. Közös jellemzőjük a TPL^h régió, ami három rövid, nem transzlált exon összeépülésével áll elő. A kódoló szakasz elhelyezkedése és terminációja alapján L₁–L₅ nevű csoportokba osztott mRNS-ekről képződnek a késői fehérjék. A IVa2 fehérje az MLP aktiválásában játszik szerepet.

A vírusrészecskék érésében kulcsszerepet játszik a virális proteáz, mely e vírusszaládon belül erősen megőrzött cisztein–proteáz. Szerepe a fent tárgyaltakon kívül az, hogy egy sor, prekursor formában kifejeződő struktúrfehérjét specifikus hasítással alakít aktív formává, hiányában a keletkező vírusrészecskék nem fertőzőképesek.

1.4. Taxonómia

Az ICTV első publikációjában⁴⁷ az adenovírusok még nemzetség szintjén szerepeltek, de 1975-ben az egységet már család szinten különítették el, amely a *Mastadenovirus*, és *Aviadenovirus* nemzetségből állt. Akkoriban a nemzetség alatti szinten a rendszertani besorolás a szerológiai vizsgálatok alapján nyugodott. A két nemzetség elkülönítése a gazdaállatok alapján történt, amit alátámasztott, hogy a két nemzetségre egyaránt jellemző antigént nem sikerült kimutatni. Figyelembe vették ezen kívül az eltérő biofizikai, biológiai és biokémiai jellemzőket is (például az aviadenovírusok viszonylag magasabb DNS- és alacsonyabb fehérjetartalmát).¹³

A faj fogalmának – ami akkoriban megegyezett a szerotípussal – definiálása is „immunológiai különbözőség” alapján történt, amint azt az ICTV második adenovírus jelentésében meghatározták⁴⁸: az egy fajba tartozó vírusok nem mutatnak neutralizációs keresztreakciót más fajokkal, vagy ha igen, akkor a homológ–heterológ titer arány 16-nál nagyobb mindkét irányban. Abban az esetben, ha ez 8 és 16 közé esik, külön fajról csak akkor beszélhetünk, ha a hemagglutininek eltérőek (azaz a hemagglutináció–neutralizációs keresztreakció negatív), vagy a DNS alapvető biofizikai, biokémiai sajátosságokban különbözik. Ez a besorolás több felmerülő problémát nehezen tud kezelni.

- A szerológiai alapon felállított alcsoportokon belül számos vírus mutat keresztreakciót típus-specifikus próbákban.

^gmajor late promoter

^htripartite leader sequence

- Régóta ismert, hogy az azonos alcsoportban lévő vírusok közötti gyakori rekombináció miatt – ha az a hexon és a fiber gének között következik be – a szérum-neutralizációs és a hemagglutináció-inhibíciós próbák eltérő típust jeleznek.⁴⁹
- Kiderült, hogy rekombináció nemzetség szinten különböző vírusok között is bekövetkezhet, amire a példát éppen egy atadenovírus és mas-tadenovírus fiber-génjeinek rekombinációja szolgáltatta.⁵⁰
- Az epitopok, amelyek alapján a felosztás történt, a szerkezeti fehérjék kis, többnyire felületi részletei csupán, tehát a teljes víusról – bár annak azonosítására többnyire alkalmasak – csak minimális információt hordoznak.
- A vírusok rendszerének filogenetikai alapokra helyezéséhez, mivel a származási viszonyokról szinte semmi információt nem nyújt, teljesen alkalmatlan ez a felosztás.

E problémák miatt nehéz volt alátámasztani bármiféle rendszertani változtatás megalapozottságát a vizsgálati módszerek változtatása és a faj definíciójának megváltoztatása nélkül. Az ICTV hetedik jelentésében⁵¹ már új species definíció jelenik meg, melyben a hangsúly a vírus DNS jellemzőinek vizsgálatára helyeződik át. Ezzel lehetővé válik az adenovírusok családjának filogenetikai alapon való felosztása. A faj szerinti besorolás az alábbiak közül legalább két jellemző figyelembe vételén alapul:

- megfelelően számított filogenetikai távolság, aminek a proteáz, a pVIII, a hexon és a DNS-polimeráz aminosav-sorrendek „distance matrix” analízisén kell alapulnia,
- DNS hibridizáció,
- restrikciós enzim fragmentum mintázat (fizikai térkép),
- a genom százalékos GC tartalma,
- daganatkeltő képesség rágcselekben,
- szaporodási jellemzők,
- kereszt-neutralizáció,
- rekombinációs képesség,
- VA RNS gének száma,

- hemagglutinációs jellemzők,
- Az E3 régió génszerkezete

A kereszt-neutralizáció hiánya és 10%-nál nagyobb filogenetikai távolság külön fajba sorolást von maga után. Amennyiben a filogenetikai távolság kevesebb, mint 5%, és a fenti jellemzők közül még bármelyik jellemző azonos csoportba sorolja a két törzset, akkor azokat azonos fajnak kell tekinteni még akkor is, ha a gazdaszervezet különböző.

1.4.1. Az adenovírusok javasolt rendszere

Az adenovírusokat a hetedik ICTV jelentésben leírt alapelvek alapján a még hivatalosan el nem fogadott rendszer szerint négy nemzetségbe soroljuk,⁵² amelyek genomszerveződési jellegzetességeit az 1.5 részben tárgyalom:

***Mastadenovirus* nemzetség**

Ide tartozik az összes korábban is ide sorolt mastadenovírus, kivéve azokat, amelyek az atadenovírusokhoz kerültek.

***Atadenovirus* nemzetség**

A Bartha Adorján által kettes altípusba¹² sorolt szarvasmarha adenovírusokon kívül az alábbi ismertebb vírusok tartoznak ide: a már említett OAV287 juh-adenovírus izolátum, a DAdV-1, egy kígyó adenovírus izolátum (Benkő Mária és Farkas Szilvia szóbeli közlése), az amerikai fekete farkú szarvasból (*Odocoileus hemionus hemionus*)⁵³ és az erszényes rókakuzuból származó vírus.

***Aviadenovirus* nemzetség**

A DAdV-1 és a TAdV-3 kivételével a madár gazdából izolált adenovírusok nemzetsége.

***Siadenovirus* nemzetség**

Ide jelenleg két ismert vírus tartozik: a leopárdbékából (*Rana pipiens*) izolált béka adenovírus egyes szerotípusa (FrAdV-1⁵⁴ⁱ, NC_002501.1) és a

ⁱfrog adenovirus 1

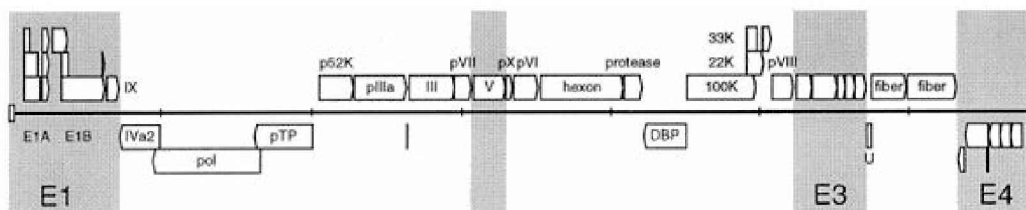
pulyka adenovírus hármasszerotípusa (TAdV-3^{20j}, NC_001958.1).

1.5. Genomszerveződés

Az alábbiakban a jelenleg érvényben lévő rendszertől eltérően, a jobb áttekinthetőség érdekében, a javasolt új felosztás szerint mutatom be az egyes adenovírus csoportok jellemzőit.

Az adenovírusok kettős szálú DNS vírusok. Az 5' végi nukleotid minden esetben citozin, melyhez mindkét szálon kovalensen vég-fehérje (TP) kapcsolódik. A genom két végének szekvenciái egyformák, ezeket rövid fordított ismétlődéseknek (ITR^k) nevezzük. A genom középső részét foglalják el az összes nemzetségben előforduló, homológ, erősen konzervált gének – a csak a mastadenovírusoknál meglévő V-ös protein génje kivétel a szabály alól. A két végen az egyes nemzetségek speciális génjei találhatók. A nemzetségenként eltérő genomszakaszok háttére sátriozott^l.

1.5.1. Mastadenovírusok



1.2. ábra. HAdV-40 géntérképe

A mastadenovírusokra jellemző az E1, E3, és E4 korai régiók megléte. Ezek a többi nemzetségben nem minden esetben fordulnak elő. Az ITR-ek viszonylag hosszúak (102–368 bp). A genom hossza 30–36 kbp körüli, GC tartalma a fajra (régábban subgenus) jellemzően változó. A IX és V szerkezeti fehérjék génjei csak ebben a nemzetségben találhatók meg (1.2 ábra).

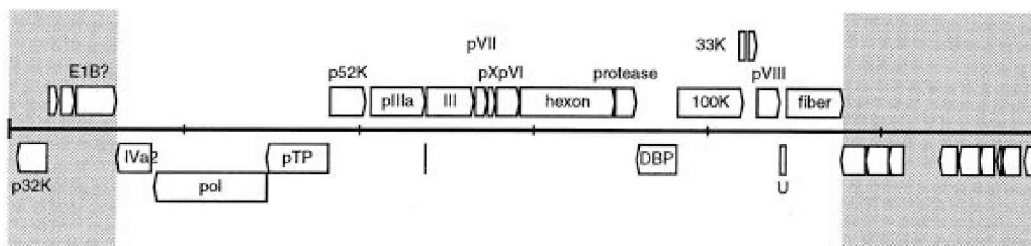
1.5.2. Atadenovírusok

Ebbe a nemzetségbe viszonylag rövid genommal (28–32 kbp) rendelkező vírusok tartoznak. Az ITR-ek rövidek (51–74 bp), az AT tartalom 50% fölötti.

^jturkey adenovirus 3

^kinverted terminal repeat

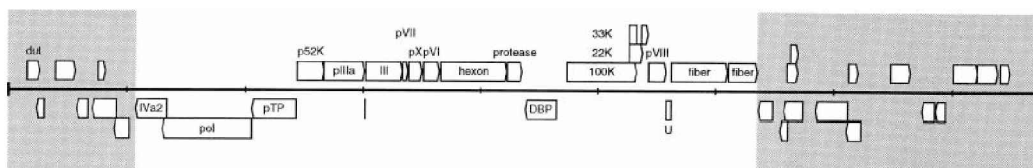
^lAz ábrák a Davison és munkatársai⁵⁴ által közölt összefoglaló ábra részei.



1.3. ábra. OAV287 géntérképe

Az E4 régió ORF-ei távoli homológiát mutatnak a mastadenovírusokéval. Az E1A gén és az E3 régió hiányzik. Az előbbi helyén egy jobbról balra átíródó struktúrális fehérje génje helyezkedik el, az utóbbival analóg szakasz pedig a genom jobb végén található. A bal végen az E1B-nek nevezett szakasz csak minimális homológiát mutat a mastadenovírusok E1B régiójával (1.3 ábra).

1.5.3. Aviadenovírusok

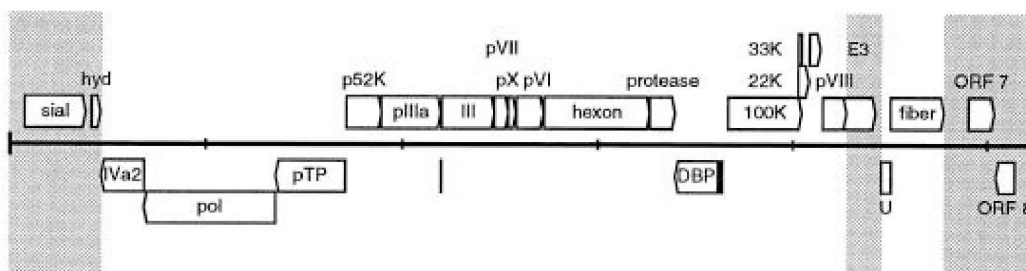


1.4. ábra. FAdV-1 géntérképe

Ebben a nemzetségben hiányzik az E1 és E3 korai régió, míg az E4 át-helyeződött, vagy hiányzik. Az ITR-ek rövidek (54–71 bp), ugyanakkor az adenovírusok a leghosszabb genommal rendelkező adenovírusok (42–44 kbp). Az erre a nemzetségre jellemző korai régiók részletes jellemzése még nem készült el. A genom jobb végének közelében van egy kis ORF-ekből álló, E4-re emlékeztető rész, ezen kívül a bal végen 2 kbp körül kezdődő ORF mutat némi hasonlóságot a HAdV-2 E4 ORF1-gyel (1.4 ábra).

1.5.4. Siadenovírusok

A nemzetség a genom bal végén található, szialidáz-enzimeket kódoló génekkel homológ génről kapta a nevét. Az ITR-ek ebben a nemzetségben a legrövidebbek (36–39 bp), csakúgy mint a teljes genomok. Az eddig ismert



1.5. ábra. FrAdV-1 géntérképe

két vírus (TAdV-3^m, FrAdV-1ⁿ) genomja mindössze 26 263, illetve 26 163 bázispár hosszúságú (1.5 ábra).

1.6. Célkitűzések

A Rus elnevezésű szarvasmarha adenovírus izolátum egyetlen irodalmi említése egy 1993-ban megjelent cikkben található, ami az EDS (DAdV-1) vírus géntérképezésével és homológia-vizsgálatával foglalkozik.⁵⁵ Abban a BAdV-7 szerotípus 1881-es törzseként írták le.

Mikor a Rus adenovírus szekvenálását 1997-ben kezdtük, az adenovírusok taxonómiájában mereven tartotta magát a két nemzetségen alapuló felosztás. Mindössze egy „atipikus mastadenovírusnak” leírt vírus genom-szekvenciája volt ismert, az OAV287 törzsé,¹⁸ ami a II. alcsoportbeli szarvasmarha adenovírusokkal mutatott rokonságot.⁵⁶ Noha az alcsoport elkülönítését Bartha¹² már húsz évvel korábban javasolta, és a szintén rokon DAdV-1 nemzetség szintű elkülönítésre is történtek próbálkozások,⁵⁷ a DNS-szekvencia adatok ismerete nélkül a genus-szintű szétválasztást a meglévő ismeretek alapján nem lehetett meggyőzően alátámasztani. Ezért a Rus vírus szekvenálásával egyik célunk annak bizonyítása volt, hogy az OAV287-nél tapasztalt genomszerveződés nem kivételes, nehezen magyarázható jelenség, hanem létezik egy jól körülírható, a korábbi két nemzetséggel viszonylag távoli rokonságban álló adenovírus-csoport, melynek a nemzetség szintű elkülönítése indokolt.

A másik cél az adatgyűjtés volt: az új nemzetségre jellemző új gének több fajból származó homológjainak ismeretében ugyanis lehetőség van az adott vírusok törzsfajlására, esetleges gazdaváltásaira vonatkozó következtetések lezárására is.⁵⁸ Erre a célra az adenovírusok, azon belül pedig az atadenoví-

^mturkey adenovirus 3 – hármas típusú pulyka adenovírus

ⁿfrog adenovirus 1 – egyes típusú béka adenovírus

rusok kiválóan megfelelnek, hiszen adenovírusok az összes gerinces osztályból ismertek, és atadenovírusok eddig ezek közül már három osztályból előkerültek (hüllők, madarak, emlősök).

Harmadrészt szándékunkban áll a későbbiekben az atadenovírusok gén-expressziójának vizsgálata, és ehhez szükségünk volt olyan célobjektumra, amely jelentős mértékben különbözik a más laboratóriumokban vizsgált rokon vírusoktól.

2. fejezet

Anyagok és módszerek

2.1. Felhasznált vegyszerek és enzimek

A felhasznált vegyszerek többségét a SIGMA, a Spektrum3D, és a GIBCO cégek magyarországi forgalmazóitól szereztük be. A restrikciós enzimek nagy része Fermentas, a többi New England Biolabs, GIBCO és QBIOGENE gyártmány volt.

2.2. Molekuláris és mikrobiológiai technikák

2.2.1. Víruszaporítás

A Moszkvai Mezőgazdasági Akadémiáról kapott „Rus” nevű szarvasmarha adenovírus törzset 10% foetális borjú savót tartalmazó MEM^a tápoldatban növesztett, primer borjú here szövettényészeten szaporítottuk. A DNS-tisztításhoz szükséges vírusrmennyiséget 2500 ml-es hengeres üvegpalackokban állítottuk elő forgatótoronyban.

2.2.2. A vírus-DNS tisztítása

A vírus részecskék tisztítása a sejtkárosító hatás (CPE) maximumakor (körülbelül 6 nappal a fertőzést követően), a sejttenyészet háromszori fagyasztása–olvasztása és a sejttörmelék centrifugálással való eltávolítása után ultracentrifugálással történt. A DNS-t proteináz K enzimmel történő emésztést és fenolos extrakciót követően etil-alkohollal csaptuk ki.⁵⁹

^amodified Eagles medium

2.2.3. PCR reakciók

A polimeráz láncreakciók körülményei megegyeztek az enzimek gyártóinak leírásával. (A tapadási hőmérsékletek 2–4°C-kal a primerek olvadáspontja alatt voltak, az elegyek térfogata 25, vagy 50 μ l volt.) Az olvadáspontok kiszámításához az egyszerű $T_m = N_{A+T} + N_{G+C}$ képletet használtuk, ahol $N_{nukleotid}$ az adott nukleotidok száma. A szekvenáláshoz sokszorozott fragmentumok esetében a Taq polimeráz által okozott esetleges másolási hibákat 4 független reakció termékének összekeverésével küszöböltük ki.

2.2.4. Molekuláris klónozás

Restrikciós emésztések

Az emésztések általában 50 μ l térfogatban az előírt mennyiségű enzim felhasználásával történtek. A puffereket a Boehringer Mannheim GmbH által javasolt készletből válogattuk.

Végfragmentumok klónozása

A *Pst*I emésztést követően az emésztményt egy órán át 37°C-on inkubáltuk 0,5M piperidin jelenlétében, majd a piperidint többszöri beszárítás segítségével távolítottuk el. A renaturálás 56°C-on történt 1M NaCl-ot és 10mM pH:7.6 Tris-Cl puffert tartalmazó közegben.

Ligálás

A vírus-DNS fragmentumok klónozásához pBluescript SK és KS „Phagemid”-eket⁶⁰ használtunk. Az enzimek hatástalanítása eleinte fenol-kloroformos extrakcióval történt, amit alkoholos kicsapás követett.⁶¹ Később áttértünk a QBIogene által forgalmazott „Miniprep Express Matrix TM” használatára. A módszer a készítményhez mellékelt, alkalikus lízissel történő plazmid DNS tisztítási leírás adaptációja volt. A kész emésztéshez az összes DNS megkötéséhez szükséges mennyiségű, de minimálisan 1 térfogatnyi mátrixot adtunk, majd az izolálást a javasolt módon fejeztük be. A ligálás minden esetben 10 μ l térfogatban, szobahőmérsékleten, 1-2 óra időtartammal történt.

DNS fragmentumok izolálása

A fragmentumokat TAE pufferben,⁶¹ agaróz gélben történő elektroforézist követően GENE CLEAN kit felhasználásával tisztítottuk.

Baktériumok transzformálása

A ligátumok kompetens *E. coli* DH5 α baktériumokba juttatása kétféle módszerrel történt:

- Gyakran elektrokompetens sejtek előállítása után elektroporációval⁶¹ végeztük.
- Az esetek más részében kémiaiilag kompetenssé tett sejteket állítottunk elő. Ehhez kétértékű kationokat (Mg²⁺, Ca²⁺) felhasználó általánosan elterjedt módszerek egy nagy hatékonyságú változatát használtuk⁶², melynek lényege, hogy a kation, ami ebben az esetben magnézium ion, az *E. coli* sejtek szaporításhoz szükséges táptalajban van. A transzformálás és a sejtek fagyasztva tárolása is ennek a táptalajnak glicerinnel és PEG-gel kiegészített változatában történik.

A kapott klónok tesztelése

Az ellenőrzéshez szükséges plazmidmennyiséget 1,5 ml baktériumkultúrából kiindulva „alkalikus lízis” módszerével állítottuk elő. Eleinte a tisztítást teljes egészében a Sambrook kézikönyvben⁶¹ leírt „alkalikus lízis” módszer szerint végeztük, később áttértünk az „Express Matrix” (QBIOSOURCE) használatára, a hozzá ajánlott módszerrel.

Szubklónozás

Erre a feladatra kétféle módszert alkalmaztunk.

- A szekvenálás, vagy a klónozás után végzett tesztelő emésztések során talált, alkalmas helyen lévő, restriktációs hasítóhelyeket használtuk fel.
- A fragmentumok izolálása után kis specifitású enzimekkel (pl. *Sau3AI*, *HaeIII*) történő részleges emésztésekkel különböző hosszúságú fragmentumokra hasítottuk az eredetit. Ezek közül azokat a fragmentumokat klónoztuk, amelyek megőrizték az izoláláskor meglévő egyik végüket. Ezt a fragmentum izolálásához használt enzimmal és a gyakran vágó enzimmal kompatibilis véget hasító, nagy specifitású enzimmal (pl. *PstI*–*BamHI* vagy *PstI*–*EcoRI*) emésztett vektorba történő ligálással értük el. Ennek a módszernek az az előnye, hogy egyetlen lépésben számos, különböző mértékben rövidült fragmentumot kapunk. A hátránya viszont az, hogy a kis specifitású enzimek jelentős része – a környezettől függően – bizonyos hasítóhelyeket előnyben részesít, így a különböző hosszúságú fragmentumok előfordulása általában nagyon egyenlőtlen.

2.2.5. DNS-szekvenálás

A kiválasztott klónok szekvenálását helyben az Országos Állategészségügyi Intézettel kooperációban az ott lévő ALF-Express automatikus szekvenáló gépen végeztük. PCR-fragmentumok közvetlen szekvenálására a Szegedi Biológiai Központ szekvenáló szolgáltatását vettük igénybe. A szükséges plazmid-DNS-t eleinte a GIBCO cég által forgalmazott „Concert Rapid Plasmid Miniprep TM” készletével tisztítottuk, később pedig áttértünk a már említett „Express Matrix” készítményre.

2.3. A DNS-szekvenciák szerkesztése, analízise és ábrázolása

Az automata DNS-szekvenáló által produkált fájlokat a „Staden alkalmazáscsomag”⁶³ „trev” programjával szerkesztettük és alakítottuk ún. kísérleti fájlá. Utóbbi célra felhasználtuk a „pregap4” programot is, majd a fragmentumok és a genom teljes szekvenciáját a „gap4” programmal készített illesztések javításával állítottuk össze.

A nyílt leolvasási kereteket (ORF^b) a csomag „spin” programjával azonosítottuk. A fehérjéket kódoló szakaszok kiválasztására a kodonpreferencia alapján való keresést használtuk. Ezt szükséges esetben a „splicing site” kereséssel egészítettük ki.

Homológiai keresése

Homológiai keresésére általában az NCBI Blast adatbázis-kereső szolgáltatását és az intézeti szerveren saját adatbázisokkal szemben futó blast2⁶⁴ programot vettük igénybe. Hasonló célra használtuk a „MEME”^{65c} rendszer „MAST”^{66d} szolgáltatását is, a „Fred Hutchinson Cancer Research Center” „Block Maker”^{67e} által képezett „homológia blokkok” segítségével.

A fehérjék általános tulajdonságainak vizsgálatára a PredictProtein szerveren^{68f} elérhető programokat használtuk.

^bopen reading frame

^cMultiple EM for Motif Elicitation

^dMotif Alignment & Search Tool <http://meme.sdsc.edu/meme>

^ehttp://blocks.fhcrc.org/blocks/blockmkr/make_blocks.html

^f<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>

Illesztések

A homológ fehérjék illesztése az esetek egy részében a Clustal W program⁶⁹ alkalmazásával történt, amihez esetenként a Clustal X^g grafikus felület segítségét vettük igénybe^h. Hasonló célra, valamint az illesztések szerkesztésére a Michele Clamp által készített „Jalview” programot használtukⁱ.

Az illesztések publikálható formáját az „alscript”⁷⁰ formázó-program segítségével hoztuk létre.

2.3.1. Filogenetikai számítások

A filogenetikai számításokhoz a PHYLIP⁷¹ programcsomag egyes alkalmazásait használtuk.

A távolság-mátrix^j alapú vizsgálatokhoz az adatsorok számítása a DNADIST („Kimura 2” paraméter algoritmus) és PROTDIST (Dayhoff PAM 001 mátrix) programokkal történt, a zárójelben megadott paraméterekkel. A távolság-mátrixokból a fákat a FITCH programmal állítottuk elő. A parszimónia^k analízist a PROTPARS programmal végeztük.

Az illesztésekből SEQBOOT programmal állítottunk elő „bootstrap” adatsorokat 100 illetve 1000 ismétlést használva.

A konszenzusos fa kiszámítására a CONSENSE programot használtuk. A fák megjelenítése a TreeView⁷² és Treetool^l programokkal történt.

Egyes fák készítéséhez felhasználtuk a PAML csomagot,^{73;74} amellyel a „maximum likelyhood”-számításokat végeztük Dayhoff⁷⁵ empirikus modellje alapján. Az számítások menete általában a következő volt:⁷⁶

1. A szekvenciák illesztése, zavaró szakaszok (deléciók, hipervariábilis régiók) eltávolítása
2. Egyes esetekben a „bootstrap” adatsorok előállítása (SEQBOOT)
3. A távolság-mátrixok kiszámítása (DNADIST, PROTDIST)
4. A fákhöz szükséges adatsorok számítása (FITCH, DNAML, DNAPARS, PROTPARS, PAML)

^gEz a program az „NCBI SOFTWARE DEVELOPEMENT TOOLKIT” része

^hncbi.nlm.nih.gov

ⁱReferencia híján az elérhetőség: <http://www.ebi.ac.uk/michele/jalview/>

^jdistance matrix

^kparsimony

^lMike Maciukenas által írt program, karbantartó: Mike McCaughey, University of Illinois

5. A konszenzus fa kiszámítása (CONSENSE)
6. Megjelenítés és a „bootstrap értékek felírása” (TreeTool vagy TreeView)

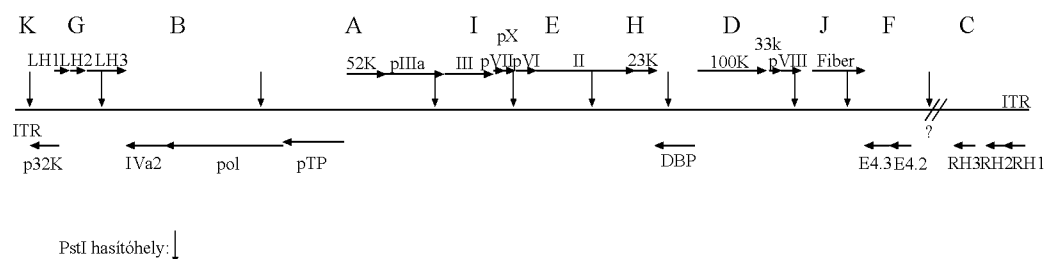
3. fejezet

Eredmények és megbeszélés

3.1. A Rus törzs genomszerveződése

Munkánk során sikerült meghatározni a Rus adenovírus izolátum szinte teljes DNS szekvenciáját. A hiány a jobb oldali végfragmentum, azaz „C” bal oldalán tapasztalható delécióból adódik. Ez feltehetőleg a végfragmentum-klónozásnál a TP eltávolításakor bekövetkező denaturációt követő rendellenes hibridizáció miatti klónozási hiba eredménye. A *Pst*I – tompa végű végfragmentum genomhoz viszonyított bal oldalának helyét a megelőző *Pst*I fragmentum bal oldala foglalja el. A hibás renaturáció valószínűleg a jobb vég közelében lévő ismétlődések miatt alakul ki, méghozzá igen nagy hatékonysággal, hiszen több ilyen hibás fragmentumot is kaptunk, deléciónélkül viszont egyet sem. A hiányzó szakasz az azt körülvevő szekvenciákra készített primerekkel felsokszorozott fragmentum mérete alapján valószínűleg mintegy 1400–1600 bázispár hosszúságú lehet, és az E3 régió hiányzó részén kívül magában foglalja az E4 jobb végét, és a köztük lévő ismeretlen szakaszt is.

A genom molekuláris klónozáshoz használt fizikai térképe, melyen a kódolt fehérjéket is feltüntettük, a 3.1 ábrán látható. A gének kódoló szakaszait a 3.1



3.1. ábra. A Rus adenovírus vázlatos géntérképe

táblázatban tüntettük fel.

A Rus adenovírus törzs szekvenciája a többi atadenovíruséhoz nagyon hasonló genomszerveződést mutat. Szintén erről tanuskodik genomjának AT tartalma is, ami 64,4%. Feltűnő, hogy az azonosított gének között nincsen a mastadenovírusokra jellemző V, IX, és E1A sem. Ezzel szemben az egész genom szerkezete hasonló az OAV287^{18;77} a DAdV-1⁵⁷ és a BAdV-4 (NC_002685) esetében tapasztalt elrendeződésre. Közös jellemzőjük a fentiekén kívül, hogy az E1B szakasz csak igen távoli hasonlóságot mutat a mastadenovírusokéval, az E4 régió nem a genom jobb végén helyezkedik el, bár a fiber gén után, de ezt követően a genom jobb végét a többi adenovírus-genusból ismeretlen régió foglalja el. Ez a szakasz eddig még nem jellemzett, egymással rokon fehérjék ORF-jeit tartalmazza. A többi mastadenovírossal vélt analógia alapján – ez a szakasz sem szükséges a szövettenyésztésben való szaporításhoz – E3 régió-nak nevezték el az OAV287 esetében. A továbbiakban én is az ott bevezetett terminológiát használom.

Az alábbiakban az egyes genomszakaszokat a hagyományos felosztást követve funkció szerinti csoportokra bontva elemezzük. Mivel munkánk egyik fontos célja volt, hogy az *Atadenovirus* nemzetség általános jellemzőit körvonalazva alátámasszuk különválasztásának jogosságát, az ebbe a csoportba tartozó vírusokra jellemző tulajdonságokat mutató régiók és fehérjék bemutatásával lényegesen részletesebben foglalkoztam, mint az egész víruscsaládra általánosan jellemző gének ismertetésével.

3.2. Az ITR szekvenciák

A különböző adenovírusok ITR szekvenciái igen változó hosszúságúak, ezért vizsgáltuk, hogy a szekvenciák összehasonlításával milyen közös cisz szabályozó elemeket találunk a nemzetségek között, illetve e végi szekvenciák milyen tulajdonságai tekinthetők az egyes nemzetségekre jellemzőnek. Ehhez az összehasonlításhoz a már ismertek mellé a most megismert Rus, valamint a BAdV-5 vírusok ITR-szekvenciáit is meghatároztuk. Az ismert szekvenciák illesztését a 3.2 mutatja be.

ORF/fehérje	kódoló szakasz (bp)	fehérje (as)	kódoló szál
p32K	1128–343	263	l
LH1	1179–1556	125	r
LH2	1553–1921	122	r
LH3 (E1B 55K	1961–3121	386	r
IVa2*	4109–3151	337	l
IVa2*	4890– ⁴⁸⁸⁷ – ⁴¹⁵⁵ –3151	336	l
DNS-polimeráz	7745–4230	1173	l
pTP*	9235–7433	604	l
pTP*	11960– ¹¹⁹⁴⁶ – ⁹²¹⁴ –7433	602	l
50/52K	9252–10292	346	r
pIIIa	10420–11953	511	r
III	11985–13337	450	r
pVII	13390–13740	116	r
pX	13762–13977	71	r
pVI	14006–14611	201	r
hexon (II)	14634–17330	910	r
proteáz	17330–17932	200	r
DBP	19079–17940	378	l
100K	19099–20973	624	r
22K*	20906–21043	45	r
33K*	20906– ²¹⁰²⁶ – ²¹¹¹³ –21342	116	r
pVIII	21374–22042	222	r
fiber (IV)	22246–23763	505	r
„E4.3 (E4–30.8K)”	24514–23834	226	l
„E4.2 (E4–25.6K)”	25151–24491	219	l
RH3	26904 (1578)–26359 (2123)	181	l
RH2	27717 (768)–27175 (1307)	179	l
RH1	28251(231)–27739 (743)	170	l

3.1. táblázat. A Rus genomon azonosított gének kódoló szakaszai. A „kódoló szál” jelölése az adenovírusoknál hagyományos (r = rightward, l = leftward) módon történt. A jobb végen kódolt gének esetében az l szálnak megfelelő távolságokat is feltüntettem zárójelben, mivel a bal végtől számított értékek még változni fognak. A „hiba” után kódolt RH1-3 géneket vízszintes vonallal választottam le. Itt az RH3 az OAV287 esetében az RH4-nek felel meg, másik két ORF számozása azonos. *:A kódoló szakaszon belül splicingot tételezünk fel.

10 20 30 40 50 60 70 80 90

H2 CATCATCAATAATATAC CTTATTTTGGATTGAAGCCAATATGATAATGA **GGGGGTGG** AGTTTgtgacgtGGCGCGGGGGCGTGGGAACGGG

H3 CTATCTATATAATATAC CTTATAGATGGAATGGTGCCAACTGTAAATGAGGTAATTTAAAAAAGTGCGCGCTGTGTGGTGATTGGCTGC

H4 CTCTCTCTATAATATAC CTTATTTTTTTTGTGTGAGTTAATATGCAAAATAGGCGTGAAAAATTTGGGGATGGGGCGCG CTGATTGGCTGT

H9 CTATCTATATAATATAC CCCACAAAGTAAACAAAGTTAATATGCAAAATGAGCTTTTGAAATTTAACGTTTCGGGGCGG AGCCAACGCT

H12 C CTATCTAATAATATAC CTTATACTGGACTAGTGCCAATATTAAAAATGAAGTGGCGCTAGTGTGTAATTTGATTGGGTGG AGGTGTGGCT

H41 CATCATCAATAATATAC CTTAAAGCTGGAAACGAGCCAATATGATAATGAGTGAGGAGGGGCTAGGGGTGG TGCGAAGtgacgtATAGGT

S8 CATCATCAATAATATAC CTTATTTTGGACTGCAAGCCAATATGATAATGAGTAGGCGGCGCGAGtgacgtAAAGGGCGG ACCCGTGAGGT

C1 CATCATCAATAATATAC AGGACAAAGAGGTGTGGCCTAAATGTTGTTTTTTTTTTTAAAAAGTTTTGTTTGTGATTGTTTTGACAAGGTTA

C2 CATCATCAATAATATAC AGGACAAAGAGGTGTGGCTTAAATTTGGGTGTTGCAAGGGCGG GGTTCATGGGACGTCAGGTTTACGGTCACG

E1 CATCATCAATAATATAC AGGACACACGGGCGATGGGGCCAAGAAAGGGGAGG AGTTGAGGCGTGGCGGGAGG CGGGGGCGG AGGCGGGGGG

P1 CATCATCAATAATATAC CGCACACTTTTATGCCCCCTTTCTGGCGTGTGATTGGCGGAGAGGGTTGGGGTGG CGGGCTGTGATTGGC

P2 CATCATCAATAATATAC CGCACACTTTTATGCCCCCTAGGGTTGTGTTCTGATTGGCCGTTAGGGTTgtgacgtAGCGTGGGAAAAtgac

P3 CATCATCAATAATATAC CGCACACTTTTATGCCCCCTTTGTGGCTGTGATTGGCGGAGAGGGTTGGGGCGG CGGGCGG TGATTGGT

P4 CAATCAATAATATAC CGCACACAAATTTCTGATTGGTGGAGGAGCGGtgacgtCAGTGGGAGG GGCAGTGTGGGAAAGGGGCGG CGCG

P5 CATCATCAATAATATAC GGAACAAAAATGCGTGatgacgtAGACGCGGTACGTAAGGGGCGG AGCCGCTGTGACCCCTGTACGGGCGCGTC

M1 CATCATCAATAATATAC AGTTAGCAAAAAATGGCGCCTTTGTTGGCTTTGTTCCAACTGTTTTTGGGCCGAGTTGGGTTTCGTTTTTCCC

TS CATCATCAATAATATAC CTGACACTTTtgacgtAAtgacgtCAGACGTAAGTTGCAAGTGCCACGTCGTCGTGGCGCTGTCTTTTGTGAC

B1 CATCATCAATAATATAC GGAACACTTTTGGCGTGATGACATtgacgtGCTGTGCGTAAGGGGGCG TGGGAAAAATGTTCAAAGGTGCGCTGG

B2 CATCATCAATAATATAC CGTGCAATTTGTGCGTGATgacgtATACAGCGTACGTAAGGGAGG AGCTTGAAGCGTGGGAAAATTTGACCTT

B3 CATCATCAATAATCTAC AGTACACTGATGGCAGCGGTCCAACGCAATCATTTTTGCCACGTCATTTATGACGCAACGCGCGAGCGT

B4 CATTCAATATATATTAAC ATGCACC **GGGGGCG** TGGCGCGGCCATTTTGAATACTCGGCGT 59

B5 CATTCAATATATATTAAC ATGCACC **GGGGGCG** TGGCGCGGCCATTTTGAAGTCTCGGCGTTTGGCGGC 67

Rus CATTCAATATATATTAAC ATGCACC **GGGGGCG** TGGCGCGGCCATTTTGAAGTGCCCGCGGTTTGGCGGCGCGA 74

O CATTCAATATATATTAAC GTTGACAGAGGGGGG CGTGTGGGTTTT 46

D1 CATGTCATTAAATAAGAC CATGCAGAAAATGCAAAATGAGGCGAAGtgacgtA 51

F1 CTTTCATCTATAATAACCTCAAAAACTAACGAGTCATAACCGACCATAAACCGCA 54

F9 CATTCAATATATATTAACCTACATGAATGACGCTGCTGTCACTCGCACTACCGCTCCCGCGCTAGATGGA 71

F10 CATCATCTTATATAAACCGCGTCTTTTGACACACTTACAACCGCGCGCGCacgt 54

T3 CAATCAATATATATACCGCATGCTTGGGAGG GGATTTCG 39

Fr1 CAATCAATATTTGTACCGCAGGTCGTGGACCATGA 36

100 110 120 130 140 150 160 170 180

H2 GCGGgtgacgtAG 103

H3 CGGGTTAACGGCTAAAGGGGCGG CGCGACCGTGGGAAAAtgacgt 136

H4 GACAGCGCGCTTCGTTAGGGGCGG GGTAACCTGATGTGTTTAAACGGTtgacgtGAATTTTG 149

H9 GATTGGACGAGAGAAGACGATGCAAAAtgacgtCACGACGACCGGTAAACGGTCGCCCGCGAGGGCGGC 158

H12 TTGGCGTGCTTGTAAGTTTGGGCGG ATGAGGAAGTGGGCGCGCGCGTGGGAGCCGGCGCGCGCGATGTA 161

H41 AGGGCGGGTGG GAAAGGGTGG AGTGAAAtgacgtTGGGGTCGAAGGTTGGGAGTGGCGCGCGGAAGtgacgt 163

S8 GAGGGGGG CGGAGCGGTgacgtGGGCGG TGGGCGG AGCGCGGGCTGGGCGTGGCGTGCTGGGGCGG GGCGGGCGTGGGAGCGGACGTgac

C1 CACCTGTTTCAAGGCGGTTTCCACGGGAAAGACCATgacgtCAATTGGGTGTTTTTGTGGACTTTGGCCCG 161

C2 CCCTGGTCAGGGTGTTCACCGGAATGTCCAGtgacgtCAAAGGCGTGGTTTTACGACAGGGCGAGTTCGCGGACTTTTGGCGCGCGC

E1 CGGGCGGGAGG C 103

P1 CGAGAGGGTGGTgacgt AGCGTGGGGAACGTgacgtCACGTGGGAAAAAtgacgtGG 147

P2 gtAGCGTGGGAAAGTGGGTGG CGTGGCAAGGGCGGAAGTTCGATGGGGCGGGGCGGGAGG AGCGGGAGG CGGGGCGGGTGG CGTGAGGG

P3 GGAGAGGGGTGTgacgtAGCGTGGGAAACGTgacgtCGCGTGGGAAAAAtgacgtG 144

P4 CGGGTGGTgacgt CACAAAGTTGCAAAATTTGTGTGTTATTGGGAGGTCGGCCACCAGATGGCACTGTTGGACAATTTTTCCGCGCGGG

P5 GGGGTTTTCCCGCGGGCGG GCGGCCCGTGGCGTCGAGGCGGCGCGCTTTCGTTTCCA 148

M1 GGG 93

TS CTTTGGACGGGCGTTTCGCTGCTCGGTTTCCAGTTTCGGGGTTCGTTCCCGAGAACGTTGAGTCATGACAGCTGAC 166

B1 GCGG AGTTTCTGGGAGG GGAGGGGCGG GGAGTGTCGCTGTGCGTGCAGCGCGGGCGGAGCGAGTGA 159

B2 TGAACGGCG 100

B3 GSGCTGTgacgtAACTGTGGGCGG AGCGCGTCGCGGAGGCGGCGCGCTGGGCGG GGCTGAGGGCGG CGGGGGCGG CGCGCGGGGCGG

190 200 210 220 230 240 250 260 270

S8 gtATG 185

C2 CCGGGTTTTTGGGCGTT 198

P2 GCGGGGCGGGGGTGA 196

P4 CGG GCAAGTTCTCTGGGGGCGG GCGCGCGCTTCTTCTGCGCCGCGGTCACGGCA 238

B3 CGGCGGGGCGG GGC 195

3.2. ábra. Kötőhelyek: Sp1: döntött, vastag; NFI: 20. pozíció után vastag; NFIII: aláhúzott; ATF: vastag, kisbetű. Rövidítések: H: human AdV (2: J01917, 3: J01960, 4: J01964, 9: J01994, 12: X73487, 41: M20300); S8: simian AdV–8 (M28838); C: canine AdV (1: Y07760, 2: U77082); E1: equine AdV–1 (K009377); P: porcine AdV (1: L43075, 2: L43076, 3: L43077, 4: L43078, 5: L43079); M1: mouse AdV–1 (M37187); TS: tree shrew AdV–1 (V01485); B: bovine AdV (1: M17112, 2,⁷⁸ 3: M17113, 4: AF238883, 5: AF238881, 10: AF238882, Rus: AF238880); O: ovine AdV (isolate OAV287, U40839); D1: duck AdV–1 (Z86066); F: fowl AdV (1: K00940, 9: AF02125310, 10;⁷⁹) T3: turkey AdV–3 (AF074946); Fr1: frog AdV–1 (AF224336).

Az ITR szekvenciák szerepe igen fontos a vírus-DNS replikációja szempontjából. A leginkább kutatott HAdV-2 esetében az ITR-t három szakaszra osztották.⁸⁰ Az 1–18. pozíció közötti szakaszon helyezkedik el a „mag-régió”^a szekvencia, más néven „A” régió, ami nélkülözhetetlen a minimális szintű replikáció megindulásához.^{81;82} Ezt követi még két „serkentő” régió („B” 19–41. és „C” 41–51. nukleotidok között),⁸⁰ melyek gazdaeredetű transzkripció faktorok kötőhelyeit foglalják magukban.^{83;84;85;86} Ezek azon kívül, hogy fontos szerepet játszanak a replikáció megfelelő szintjének fenntartásában,^{87;88} a genom végeihez közeli promoterek „upstream” elemeiként is szolgálnak,⁸⁶ tehát a korai régiók transzkripciójának szabályozásához is szükségesek.

A 3.2 ábrából első látásra is megállapítható, hogy az ITR szekvenciák hosszuk és szerveződésük alapján a javasolt négy nemzetségnek megfelelően oszthatók fel. A mastadenovírusok szekvenciái a leghosszabbak, a siadenovírusokéi a legrövidebbek. Az atadenovírusok ITR szekvenciái nagyjából az aviadenovírusokéhoz hasonló hosszúságúak, összetételük viszont azokétól eltérő.

Megfigyelhető az is, hogy a mag-régió szekvenciáján kívül teljesen megőrzött nukleotid még nemzetségen belül sem található, bár vannak közeli rokon csoportok, melyeknél ez a szakasz teljesen egyforma. A víruscsalád minden eddig vizsgált tagjában teljes konzerváltság mindössze két helyen, az 1. és 17. citozin esetében figyelhető meg. Ezen kívül a 17. hely előtti 8–11 nukleotid hosszúságú, A,T összetételű szakasz konzervált, amit egy – az EDS kivételével – G mentes 4–6 nukleotid hosszúságú szakasz előz meg.

A mag-régió az AT összetételű része az aviadenovírusoknál szembevetően a legrövidebb, ami talán a gazdaállatok magasabb testhőmérsékletével állhat összefüggésben. Erre utal az is, hogy a DAdV-1-nél is hasonló hosszúságot találunk. (A madár eredetű adenovírusok DNS-ének AT tartalma általában viszonylag alacsony.) A leopárdbéka (*Rana pipiens*) adenovírusának genomjában látható rövid AT-ben gazdag szakasz ennek ellentmondani látszik, de ha figyelembe vesszük, hogy észak-amerikai hideg vizű patakokban és tavakban élő kétéltűről van szó, akkor láthatjuk, hogy ilyen alacsony hőmérséklet mellett néhány nukleotidnyi változás által könnyített láncszétválásnak nemigen lehet szerepe.

Az a korábbi feltevés, miszerint az ITR szekvenciák „magja” a „TAAT” szekvenciárészlet, ami az adenovírusok törzsfejlődésének korai szakaszában konzerválódott,^{89;90} feltehetőleg csak a *Mastadenovirus* nemzetségre igaz.

A mag-szekvencia 9–18. pozícióján végzett mutagenizációs vizsgálat⁹¹ során a fentiekkel összhangban kísérletesen is igazolták, hogy a 12–16. hely kö-

^acore-origin

zött egyes bázisok cseréi nincsenek hatással az *in vitro* DNS-replikációra. Az is érthető, hogy a 17. és 18. citozin cseréjénél (HAdV-2) miért tapasztalható a replikáció megszűnése.⁸⁰ Tehát a szakasz általános összetétele a fontos; valószínűleg a DNS másodlagos szerkezetének és/vagy könnyű „olvadásának” fenntartása miatt.

E mutagenizációs vizsgálatból azt is megállapították, hogy a mag-régió kívüli „A” és „B” szakaszok nemcsak hogy szükségesek a normális replikáció fenntartásához, hanem elhelyezkedésük is rögzített, továbbá a folyamatban fontos szerepet játszik az NFI⁸⁰ gazdafaktor is. Ezekről a következtetésekről a 3.2 ábrán látható illesztés alapján látszik, hogy legfeljebb a humán adenovírusok egy részére igazak, hiszen a HAdV-9-nél az NFI^b kötőhely máshol található, és a konszenzustól (TTGGCNNNNNGCCAA) erősen eltér. A HAdV-4-nél, valamint a sertés-adenovírusoknál nem is találtuk meg, nem beszélve a többi nemzetségről, ahol az ITR szekvenciák a mag-régió kivételével teljesen más szerkezetűek. Ezeken az „A” és „B” szakaszok egyáltalán nem találhatók meg.

Említésre érdemes még az tény, hogy a szarvasmarha-eredetű atadenovírusoknál az ITR-en az NFIII^c és ATF gazdafaktorok kötőhelyei hiányoznak, az NFI kötőhelyek pedig részben átfednek az Sp1 kötőhelyekkel, ami összefügghet e vírusok gyenge szaporodó-képességével.

3.3. Korai régiók

3.3.1. A bal oldali korai régió

Az „LH” fehérjék

A genom bal végén, ami megfelel az OAV287 LH^d régiójának, három, start kodonnal is rendelkező lehetséges nyílt leolvasási keretet sikerült azonosítani,

Az LH1 ORF eddig csak az OAV287 (Q9QNN2), BAdV-4 (Q997I8) és Rus vírusokból ismert, az EDS vírusból az alábbiakban az „X” ORF-nél ismertetett módszerrel sem tudtuk kimutatni. Hasonló fehérjét az adatbázisokban sem sikerült találni sem egyszerű blast kereséssel, sem a p32K fehérjénél ismertetett „blokk-kereső módszerrel”. Az LH1 szakaszon kódolt fehérje létezése még kérdéses, de mivel transzkripciója részben bizonyított az OAV287

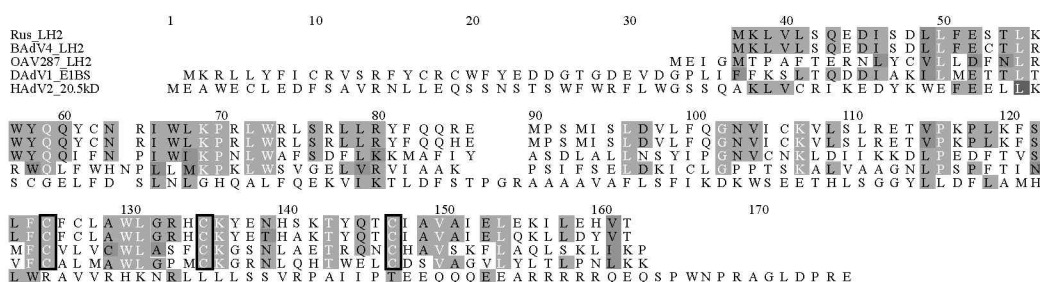
^bnuclear factor I

^cnuclear factor III

^dAz atadenovírusoknál használatos elnevezés, a „left hand” rövidítése

esetében,⁷⁷ és mivel nem fed át más génekkel, nagy valószínűséggel kifejeződik, bár lehetséges, hogy más fehérjék exonjaként.

Az LH2 fehérje – amit korábban a DAdV-1-nél⁵⁷ a HAdV-2 E1B-1 génjéről átirított 20,5 kDa proteinnel mutatott marginális homológia alapján vélték az E1B régió részének – valójában nem áll rokonságban a HAdV-2 proteinnel. Ez látható a 3.3 ábra illesztéséből is. Az atadenovírusok hasonló pozícióban talált LH2 fehérjéi viszont egymással homológok. Bár a szekvenciák meglehetősen eltérőek, egyes szerkezeti szempontból fontos aminosavak teljes mértékben megőrződtek, ezeket az ábrán keretezéssel jelöltem.



3.3. ábra. Az LH2 proteinek és a HAdV-2 20,5K fehérje illesztési kísérlete. A hasonló szakaszok árnyékoltak.

Az LH3 régió által kódolt fehérje családon belüli homológ volta, illetve az egyes nemzetségeken belüli megőrzöttsége vita tárgyát képezi az irodalomban, de feltételezték, hogy az atadenovírusok e fehérjéje a mastadenovírusok E1B 55K proteinjével rokon. Ezért elkészítettük az ismert atadenovírus fehérjék, és néhány mastadenovírus fehérje illesztését. Bár az atadenovírusoknál található változatok ebben az esetben is mintegy 100 aminosavval rövidebbek, az illesztésből egyértelműnek látszik az LH3 és E1B 55K fehérjék homológ volta (3.4 ábra). A HAdV-5 esetében végzett mutációs vizsgálat⁹² során azonosítottak olyan pozíciókat, amelyek változtatása az E1B 55K és E4 34K közötti komplex létrejöttét akadályozza. Ezek közül az egyik, az LH3 proteinekben is konzervált pozíció. Hasonlóan konzervált az mRNS-ek kötésében szerepet játszó egyik szekvenciariészlet is, ami arra utal, hogy a két nemzetség e homológ fehérjéi között funkcionális hasonlóság is van.

Ezekon kívül van a régióban még egy, a DNS polimerázzal átfedésben, azaz ellenkező irányba futó, általunk X-nek nevezett „exon” is. Erre több más adenovírus genomjának szekvenálásakor is felhívták a figyelmet: A TAdV-3 esetében²⁰ (Q9YUR5) rövid ORF-ként, a HAdV-7 esetében „hypothetical 11.5 kDa early protein” (P03288) néven írták le⁹³ a HAdV-2-nél pedig „hypothetical 11.6 kDa early protein” néven szerepel.⁹⁴ (P03287) Meglepő volt, hogy három nemzetségben is megtalálható egy ismeretlen szerepű, rövid ORF, ami igen

jó megőrzöttséget mutat. Elhatároztuk, hogy megvizsgáljuk előfordulásának általánosságát. Ehhez az NCBI Blast szolgáltatása által talált, fent említett, feltételezett fehérjékből Blast adatbázist generáltunk, és az intézetünkben is futó programmal ebben az adatbázisban kerestünk a génbankban elérhető teljes adenovírus-genomokban homológiákat. Kiderült, hogy a „fehérjének” minden adenovírusban van „homológja”, de gyakran azonos leolvasási keretben több stop kodon is található benne (a FrAdV-1 esetében például három is). Emiatt valószínű, hogy nem kifejeződő szakasról van szó. Látszólagos megőrzöttsége annak köszönhető, hogy a DNS-polimeráz konzervált szakaszával fed át, úgy, hogy a DNS két szálán kodonjaiknak harmadik, azaz legvariábilisabb pozíciói egybeesnek, vagyis a polimeráz konzerváltsága a másik szál kodonjait is erősen konzerválhatja.

A IVa2 protein

A mastadenovírusok homológ proteinjénél néhányszor sikerült kimutatni a IVa2 fehérje mRNS-e esetében splicingot is. Ilyen előfordulását az OAV287 homológ génjéről átírt mRNS-eken végzett vizsgálat nem igazolta, de valószínűleg ki sem zárta.⁹⁵ A Rus adenovírus esetében valószínűsíthető, hogy a kezdő metionint kódoló triplete splicing útján kerül az ORF elejére. Ezt a megállapítást arra alapozzuk, hogy a számítógépes analízis már jóval a splicing nélkül elérhető start-kodon előtt is kódoló szekvenciát jelez. A spin program segítségével sikerült találni egymástól igen nagy távolságban lévő, alacsony megbízhatóságú donor és akceptor helyeket, melyekkel a kódoló szakasz N-terminális irányban kissé „megnyújtható” (3.5 ábra és 3.1 táblázat).

3.3.2. Az „E3” régió fehérjéi

Az „E3” régióba a genom jobb végén lévő, sorozatos duplikációkkal kialakult szakaszon kódolt, az atadenovírus nemzetségre jellemző fehérjék tartoznak. A mastadenovírusok E3 fehérjéivel – melyek valószínűleg nem is feltétlenül mind korai fehérjék³⁸ – nem mutatnak rokonságot, és funkciójuk sem ismert. A Rus vírus ismert szekvenciáin három nyílt leolvasási keretet sikerült azonosítani a jobb vég mintegy 2,5 kilobázisnyi darabján. Ezekre az atadenovírusoknál jelenleg használt terminológia szerint RH1, RH2, RH3 néven hivatkozunk (3.1). A feltételezett fehérjékből, és más atadenovírusokban talált homológjaikból készített illesztés a 3.6 ábrán látható. Az E3 régió RH proteinjeinek leszármazási sorrendjét a rajtuk végzett filogenetikai vizsgálatok alapján szerettük volna tisztázni.

A Staden csomag „spin” programjával megvizsgáltuk, hogy DNS szinten

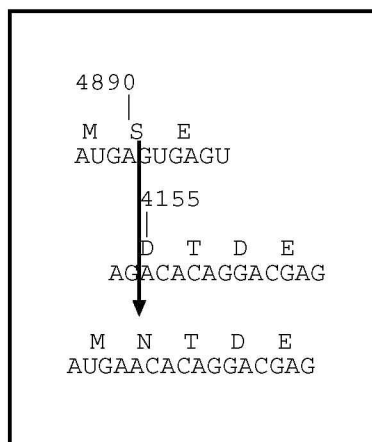
vajon milyen duplikációk nyomait lehet felfedezni a genom jobb végéhez közeli szakaszon. A 3.7 ábrán – ami tulajdonképpen a szekvencián lévő hasonlóságok diagramja – az E3 genomszakasz létrejöttében fontos szerepet játszott duplikációk nyoma jól látható. A 170–710 bp (a) közti mintegy 540 bázispárnnyi szakasz ismétlődik az 1560–2100 bp közti szakaszon (b), a 230–300 bp (c) 950–1020 bp-nál (d), a 450–530 bp (e) 1160–1240 bp-nál (f), az 1120–1260 bp (g) közti szakasz pedig 1780–1920 bp (h) között. Ha az ablakméretből és a grafikonból következő pontatlanságokat figyelembe vesszük, akkor látható, hogy (c) megfelelője a középső szakaszon (d), (e) megfelelője pedig (f) és (g), utóbbinak még kópiája viszont (h) a fragmentum jobb oldalán. Érdekes, hogy (e) és (h) nem hasonlít egymásra annyira, hogy a program jelezné. Mivel a (c)–(e) távolság megegyezik a (d)–(g) távolsággal, valószínű, hogy az (a) és (b) szakaszokkal határolt régió volt e két szakasz közös elődje, ami duplikálódott, majd ezt követően jóval később történt még egy duplikáció, aminek az eredménye (a) vagy (b) létrejött. Ha a fentiek sematikus ábrázolását (3.8 ábra) összevetjük az RH 1-3 ORF-ekkel, akkor láthatjuk, hogy a rövid, DNS-szinten megőrződött szakaszok a fehérjék „mag” régiójának, és egy kevésbé megőrzött, „A”-nak nevezett részletet kódoló szakasznak felelnek meg.

A javasolt nemzetség „E3” fehérjéinek „mag” szekvenciáján végzett „maximum likelihood” elvű filogenetikai számítás alapján megállapítható, hogy a „mag” régiók többsége egymáshoz igen közeli rokonságot mutat, vagy legalábbis egy fejlődési vonal mentén helyezhető el. Kivételt csak az OAV287 RH1 feltételezett fehérjéje képez, amely úgy tűnik, valamiféle eltérő fejlődési irányban indult el. Feltehetőleg olyan új funkciót szerzett, ami a „mag” változásával járt együtt. Ennek a feltevésnek a vizsgálata további kísérleteket igényel. A filogenetikai számítás eredményét a 3.9 ábrán tüntetük fel.

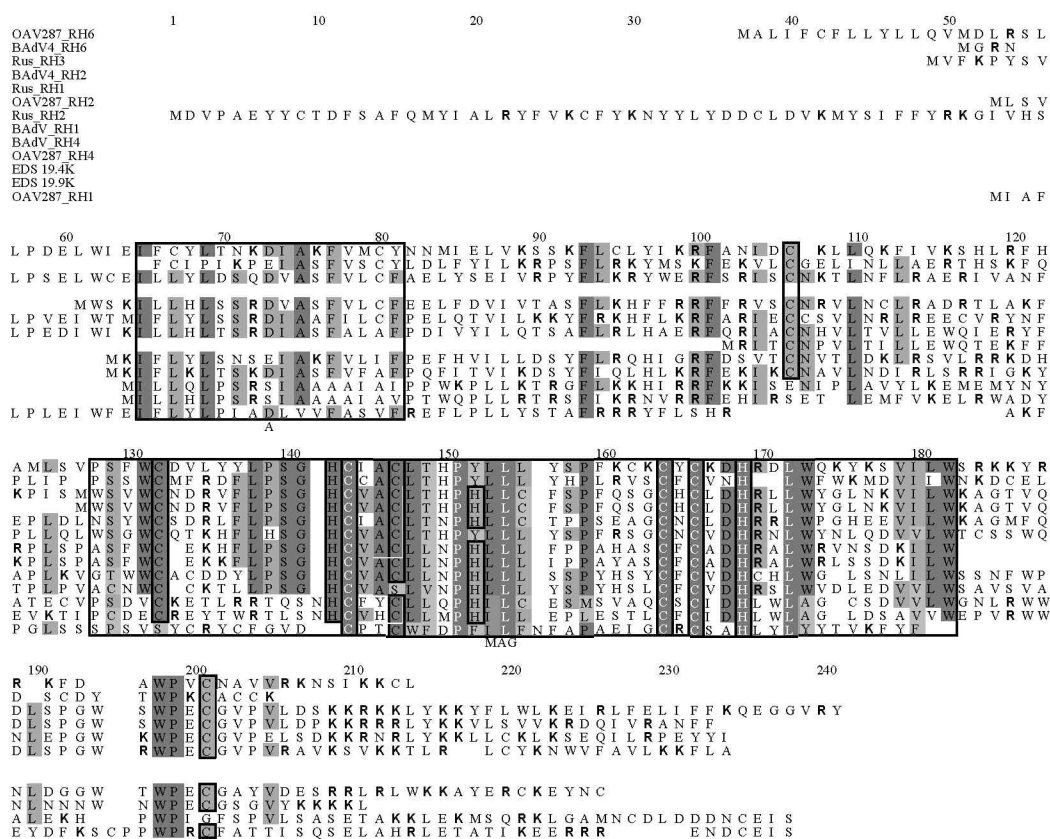
3.3.3. Az E4 régió fehérjéi

Ez a régió egyes fehérjéinek a mastadenovírusok E4orf6 vagy másnéven E4 34k fehérjéjével mutatott homológiája miatt, és mert a fiber után helyezkedik el, az E4 nevet kapta.^{57;18;77} Ezen a szakaszon két ORF-et tudtunk azonosítani, amelyeket az OAV287-nél leírt megfelelőik alapján neveztük el. A Blast keresésben hasonlóknak bizonyult fehérjék szekvenciáinak felhasználásával készített illesztés a 3.10 ábrán látható. Az E4 34K fehérje a cink-kötő fehérjék közé tartozik. A fertőzött sejtben részt vesz a késői gének kifejeződésének szabályozásában poszttranszkripció szinten, a virális DNS szintézisében, valamint a p53 aktivitásának, sejten belüli eloszlásának és lebontásának szabályozásában.^{96;97;98;99} E folyamatok jelentős részében hatását az E1B 55K proteinnel kölcsönhatásban fejt ki. Nem véletlen tehát, hogy mindkét fehérje

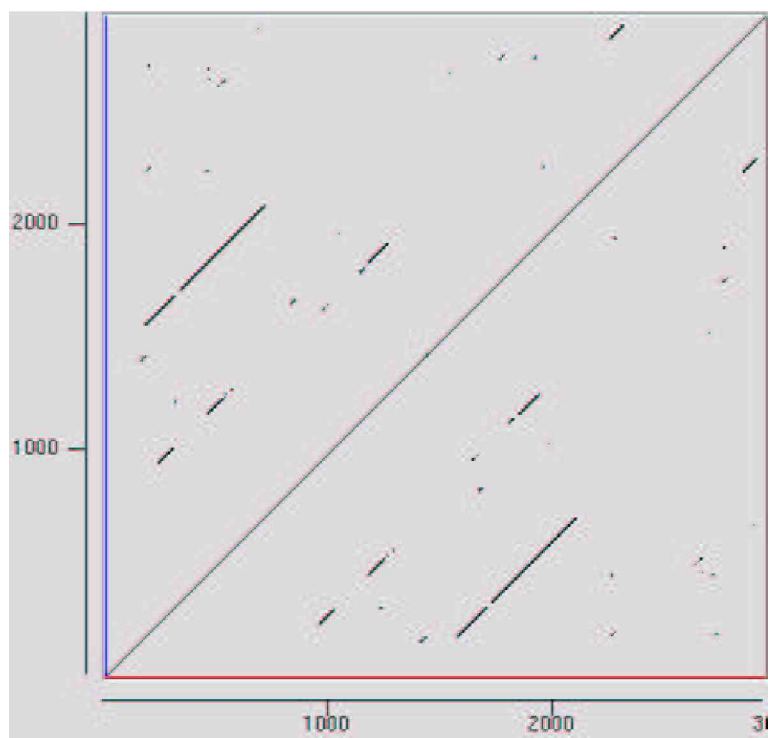
előfordul az *Atadenovirus* nemzetségben is. A HAdV-5 E4 34K proteinjén végzett mutációs vizsgálatban azonosítottak⁹⁶ néhány, a E1B 55K-val való komplexképzésben fontos szerepet játszó pozíciót. Ezek többsége a mindkét nemzetségben megőrzött rövid szakaszra esik (3.10 ábra). Az E1B 55K oldaláról is történt mutációs vizsgálat,⁹² és a komplex-képződést gátló mutációk egyike abban az esetben is konzervált pozícióra esik. Nagyon valószínű tehát, hogy a távoli homológiát mutató E4 34K E4.3 és E4.2 fehérjék a két vírustípus esetében hasonló módon kapcsolódnak az E1B 55K illetve LH3 fehérjékhez.



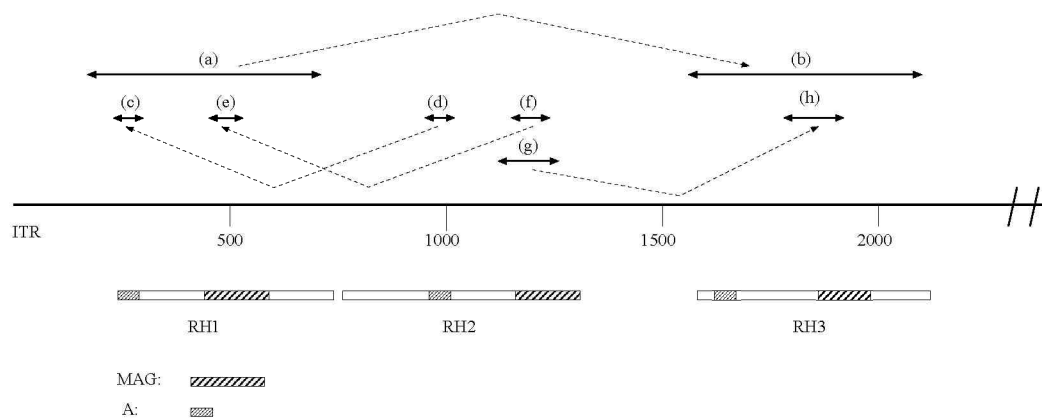
3.5. ábra. A IVa2 mRNS-ének kódoló szakaszán feltételezett splicing.



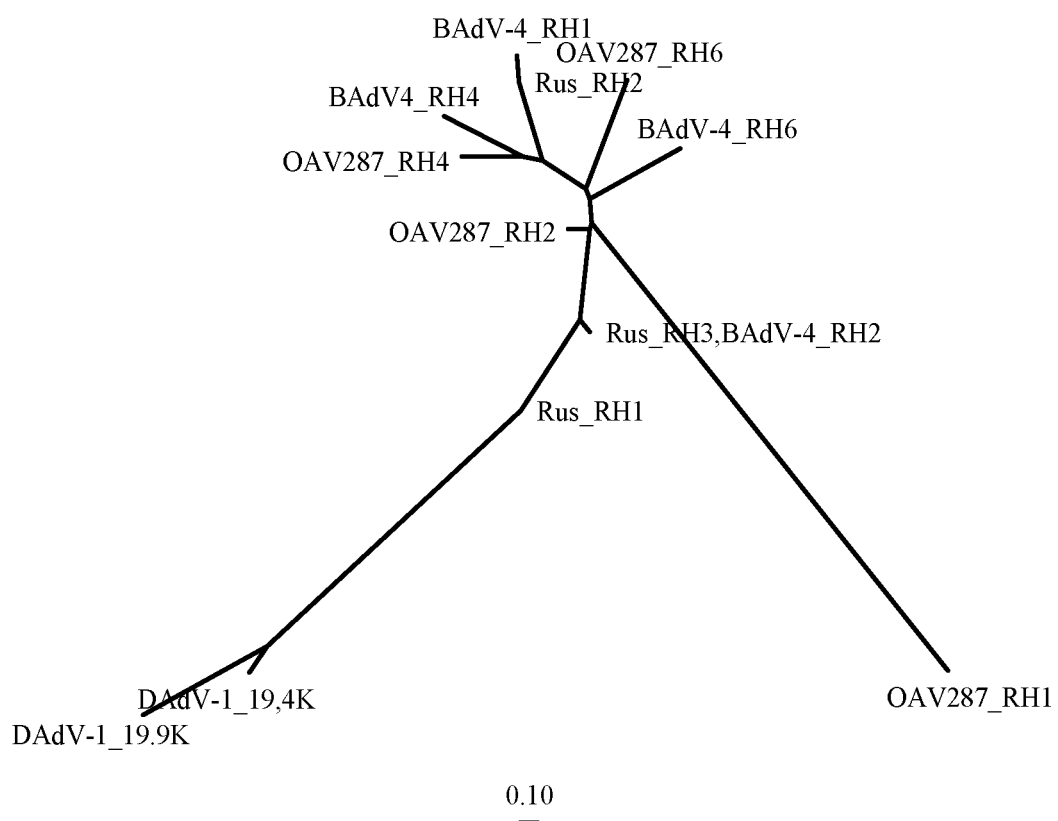
3.6. ábra. Az atadenovírusok azon E3 fehérjéinek illesztése, amelyek homológiát mutattak a Rus vírusból meghatározott ORF-ekkel. A bázikus aminosavak vastagon szedettek.



3.7. ábra. DNS szakasz ismétlődések a Rus vírus genom jobb végén. A spin program által a végfragmentumon 61 nukleotid hosszúságú ablakkal, 33 minimális egyezéssel kimutatható homológiák. A genom jobb végétől, ebben az ábrázolásban a bal alsó saroktól indul az „r” szál a két tengelyen.



3.8. ábra. Az E3 régió homológnak talált szakaszai. A spin program által hasonlóknak mutatott szakaszokat szaggatott vonallal kötöttük össze. A számozás a jobb genomvégtől indul ki.



3.9. ábra. Az „E3” szakasz fehérjéinek „mag” régiójából „maximum likelihood” módszerrel számított törzsfa. A számítás a PAML csomag programjaival történt.

1 MTTSGVPFGMTLRPTRSRRLSRRTTPYSRDRLPSPFETETRASILEDHPLLPECDNTLTMT
 HAdV-2
 HAIV-9 MQT EIQSSSLRHHPYRRARLPSPFDEETRASLMFHNCRMEGSCNAETT
 CAdV-1
 CAdV-2
 PAdV-5 MEEWVIPPITTYGYECHNLRLG
 Rus E4.2
 BAdV-4 E4.2 MELNMQLIAC
 OAV287 E4.2 MELNIQLIAC
 BAdV-4 E4.3 MDKQIKVVAC
 Rus E4.3
 OAV287 E4.3 MNLVGI
 DAdV-1 E4.3
 *

60 70 80 90 100 110 120
 HNVSYYVRGLPCSAAGFTLLIQEEFVVPWDMVLTREELVILRKCMHVCLCCANIDIMTSMMIHGYE
 HNVSYYVRGLPCSAAGFTLLIQEEFVVPWDMVLTREELVILRKCMHVCLCCANIDIMTSMMIHGYE
 SHVTAVVVRAPPIFCNCFFALCLCEIPIILWDDVLTREELVILRKCMHVCLCCANIDIMTSMMIHGYE
 CHVTAVVVRAPPIFCNCFFALCLCEIPIILWDDVLTREELVILRKCMHVCLCCANIDIMTSMMIHGYE
 NHGTDFTLSPSAVCFAMGLFFPVPVKOILTRHEELVILRKCMHVCLCCANIDIMTSMMIHGYE
 DKPHITNSHGMVEVFRCHGRGRGLCPALRVSCLLKKMLTPVIYDOYFCYFPVVSQKRDSFFKQSDMDY
 DKPHITNSHGMVEVFRCHGRGRGLCPALRVSCLLKKMLTPVIYDOYFCYFPVVSQKRDSFFKQSDMDY
 EKPSVPSHGLEIFFTCVGRGRGLGSAIRVYNLKKILTPVIYDOYFCYFPVVSQKRDSFFKQSDMDY
 KMFVNSTCYCEDSHFLLLRVDIPFFFSINNVIQSLCDHIIEFKLIYTHLCSLPAAYCQ
 CMMFVNSTCYCEDSHFLLLRVDIPFFFSINNVIQSLCDHIIEFKLIYTHLCSLPAAYCQ
 MNSIMCEGIIHHLFYCDIPFFFSINNVIQSLCDHIIEFKLIYTHLCSLPAAYCQ
 MSTIRSTCEGGHIVLWYPSGFVSIINVKELECAQLPHPTFRIMEMQLCPLP CQGF

130 140 150 160 170 180
 S W A L H C H C S S P G S L Q C L A G G Q V L A S W F R M V V
 R W L H C H C S S P G S L Q C L A G G Q V L A S W F R R V I
 M W Q V H C H C S S P G S L Q C L A G G Q V L A S W F R E F V
 T W R V H C H C S S P G S L Q C L A G G Q V L A S W F R D F I
 S W E L H C H C S S P G S L Q C L A G G Q V L A S W F R E V L
 K F F Y H C H C S S P G S L Q C L A A K Q V L L K C F D N V P
 K F F Y H C H C S S P G S L Q C L A A K Q V L L K C F D N V P
 M Y K I H C H C T D P T S L Q C L C V K Q L L L D K L K H M P
 M Y K I H C H C T D P T S L Q C L C V K Q L L L D K L K N V K
 Q N G M L H C H C D N P S L Q C Q S L K Q T L M L I L K K V H
 R G M L H C H C D N P S L Q C Q S L K Q T L M L I L K K V H
 * * * * * R A M R R I L L H V I R G N C T F R T P D R I P Y T H C C T A F T I S K T N D V V F L S S S N D V K

190 200 210 220 230 240 250
 D G A M F N Q R F I W Y R E V V N Y N M P K E I M Y M G S V F M R G R H L I Y L R L W Y D G H V G S V V P A M S F G Y
 Y G C M F N Q R F I W Y R E V V N Y N M P K E I M Y M G S V F M R G R H L I Y L R L W Y D G H V G S I I P N M S F G W
 M G G F V N K K Y L W Y R E F V N S S R P D E I N Y V G S I I F R N T H Y I Y F R L S F F R T V H Q A C M L A I Q R C I S P E L
 K G G A M N K K Y M W Y R E F V N S S R P D E I N Y V G S I I F R N T H Y I Y F R L S F F R T V H Q A C M L A I Q R C I S P E L
 N G T C Y N N M F W E Y R E Y V N R G M P E R L L Y T G S V M F R G I H L I F I R L L D S D S I R I R K L N C G A
 V N L F S E R V L R T V P V C C N T W V C C P I I F L G A F N M K Q K H C Y F L L C G P N G
 V N L F S E R V L R T V P V C C N T W V C C P I I F L G A F N M K Q K H C Y F L L C G P N G
 V T L F S E R V N K T V P V C C N T W I C C P L M F L G A S Q M H K T H F Y F L F S G V M E
 R K I P F V L P W T N C C S N V I F Y Q V V N E Y I V F C R F W N A L H V E E E I V K S H S Y
 R K I P F V L P W T N C C S N V I F Y Q V V N E Y I V F C R F W N A L H V E E E I V K S H S Y
 V R V P F V L P W T N C C T N V L C Y S F D D T V L I T R L W N A L P L E E E I S K I V T D Y
 I G D F L D H I E V P Y T M Y N A M R R R K V Y E I G R K L S V P E L N E I A Y L H R H I N L E H S G R L G Y H L G D L D T L R T

260 270 280 290 300 310 320
 S A L H C G I L N N I V V L C C S Y C A D L S E I R V R C C A R R T R R L M L R A V R I I A B E T T A M L Y S C R T R R R Q Q
 S A L N Y G L N M M Y I M C C S Y C A D L S E I R V R C C A R R T R R L M L R A V R I I A B E T T A M L Y S C R T R R R Q Q
 G V V F K S T Y N Y W L V L K C K S C S L Q N Y C A L K S C A F W V R S S I I D R V L R E V E K I P V V L H R T T S K A E E R R Q T
 G V V L K S T Y N Y W L V L K C K S C S L Q N Y C A L K S C A F W V R S S I I D R V L R E V E K I P V V L H R T T S K A E E R R Q T
 C V F Y E G M Y C N Y L I L M C K S C K H M S E I O A R V C A R R T R R L M L R A V R I I A B E T T A M L Y S C R T R R R Q Q
 Q S F K N E W F N Y Q Q L K E I T F C D S C S M L V V R C K C K F S M K E C I R R T E V R Y N F V R M L Y S C T F L F R P S N
 D D F S S K W Y S Y Q T V K R I T F S E T C E L I S I K C K C S D T L Q Q C I S R F E K R F N F L R M L Y C C S L I Y R G L S
 R R I N F G N F K S I F I I R E N I V K L R E T L A R L H I Q L N L P D L R G F H L P P Y D N C Y W K S A E T A D L Q N Q P
 R R I N F G N F K S I F I I R E N I V K L R E T L A R L H I Q L N L P D L R G F H L P P Y D N C Y W K S A E T A D L Q N Q P
 N R Y N C P M K V L F V F K S P E V Q D C K N L D T V H L K M N M V D L R G F H L P P Y D Y A Y W K S Y A S I D S N D Q P
 L T I D D A F Q N L C M L R F L S L R T C Y Q E R F I K T F L D Q P R R T H D V T P L L L P P P I D L G N D T W S C R Y A F D D C S

330 340 350 360
 F I R A L L Q H H R P I L M H D Y D S T P M
 L L R A L M E R H R P I L F S E Y E S V R S S H S T R L
 A L K Q A M M Y G R C R H I Q N L C L V N L N A F L H F
 A L R Q A M M H G R C R Q I Q N L C L V N L N A F L H F
 E L L R V M K Y G K N V S K S F I K F L
 E S G W H L
 E S G W H L
 E P S L D V A
 C L I R F V K F N G M E Y Q E R Y L K R L L K N R L R
 C L I R F V K F N G M E Y Q E R Y L K R L L K N R L R
 C L L R F V N Y R G T L C Q E R Y V K R L L F N K E F L K Y S L S E L M C C K H T N W Y I F S
 C Q R R C E C V N Y I A V T V T K E T A M K Y L D G R L E R P F A L R D Y I K L Y E F A F K

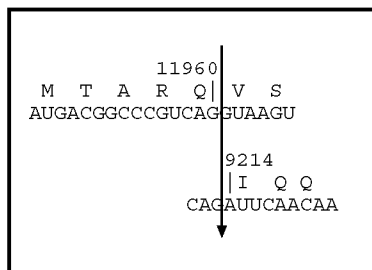
3.10. ábra. Az E4 34K fehérjék illesztése a SWALL adatbázison végzett blastp keresés eredményéből. Az ábrán feltüntettem az E4.2 fehérjéket is, mivel ezek az E4.3 közeli homológjai. Az utolsó sorokban a HAdV-2 E4 34k-E1B 55K kapcsolatot megszüntető mutációk csillaggal jelöltek.

3.3.4. A DNS–replikáció vírus eredetű faktorai

Ebbe a kategóriába az E2 régió génjei tartoznak, vagyis a DNS-polimeráz (pol), a DNS-kötő fehérje (DBP), és a terminális fehérje prekursor (pTP).

A Rus DNS-polimeráza a DAdV–1, BAdV–4 és OAV287 vírusoknál azonosítottához igen hasonló szekvenciájú. A többi atadenovírushoz hasonlóan a kódoló szakasz eleje 18 nukleotid hosszúságban átfed a pTP ORF végével.

A pTP transzkripciójának eddig vizsgált eseteiben megállapították, hogy a kódoló szakasz két darabból splicing révén épül össze: a III-as protein és a pIIIa közé eső donor-helytől néhány kodonnal jobbra lévő starthelyről indul a transláció.¹⁸ Bár pTP kódoló szakasza esetünkben az akceptor helyhez közeli startkodonnal is felírható, számítógépes analízissel mi is hasonló következtetésre jutottunk. Az eredmény a 3.11 ábrán látható. Az OAV287 esetében azt tapasztalták, hogy a pTP és az 50/52K gének kódoló szakaszai jelentős mértékben átfednek, és így a mastadenovírusoknál megszokott helyen a VA RNS géneknek nem marad hely. A Rus esetében az ORF-ek, ugyan nem fednek át, de a közöttük lévő rövid szakaszon szintén nincsen hely a VA RNS gének számára. Megjegyzendő, hogy a VA RNS gének a mastadenovírusok közül az emberi, és egyes majom vírus típusokban, a DAdV–1-ben, valamint az *Aviadenovirus* nemzetségben fordulnak elő.



3.11. ábra. A pTP mRNS kódoló szakaszán feltételezett splicing

A DBP az atadenovírusokra jellemző módon jóval rövidebb, mint a mastadenovírusnál előforduló megfelelője. (A HAdV–2-höz képest a hosszbeli különbség például 151 aminosav!). A rövidülés a variábilis N–terminális részre esik, amit a konzervált C–terminális követ. E fehérjék végeinek felépítése nemzetségre jellemző különbségeket mutat. Az atadenovírusok sajátossága, hogy a vég közelében egy NLS van, és egy glutaminsav és aszparaginsav maradékokból álló szakasz zárja a peptidláncot, míg a többi adenovírusban a véghez közel egy, vagy két fenilalanin található (3.12 ábra).

A DBP cink kötő fehérje.¹⁰⁰ A Rus esetében, ahogy a többi hasonló vírusnál is, a mastadenovírusokkal homológ H-x-C-x(8)-C-x-H motívum első

	1	10	20	
BAdV-2 AAB16762.1	P Q Y R V Q	N T L L P T G Q D D A	D E C L F	
PAdV-5 Q994D7	P Q Y Q V Q	N T I L P T G Q H D D	D D C L F	
PAdV-3 CAB41032.1	P Q F Q V Q	N T I F P V G T E D D	D E S L F	
BAdV-10 AAC35886	T Q Y Q Y Q	N T I L P V A Y V D T	D E S L F	
BAdV-10 O71102	P Q Y Q H Q	N V I L P I D Q Y D A	D D S L F	
CAdV-1 AAB05445.1	D E H R Y Q	T T I L P Q G Q H D D	D L V L F	
HAdV-2 NP_040527.1	T K H Q Y R	N V S L P V A H S D A	R Q N P F	D F
HAdV-5 P03265	T K H Q Y R	N V S L P V A H S D A	R Q N P F	D F
SAdV-25 Q8UY78	A K Y Q F R	N V S L P A G H A E T	R K N P X	D F
HAdV-4 P06500	A K Y Q F R	N V S L P A G H A E T S	R Q N P F	D F
HAdV-7 P04497	N R L Q F R	N V S L P A G H Y D S	R Q N P F	D F
HAdV-12 P04498	Q R L Q Y R	N I S L P T N H G D C	R E E P F	D F
HAdV-41 AAA52196.1	S R L Q Y R	N V A L P A S H G D G	E K E P F	
FAdV-1 AAC54914.1	E S Y A F K T	E V L A P V S P I A S	D D P F A	
FAdV-8 Q9QM71	E R E G F K T	E V L A P V K P V E T	G D P F A	
Rus	P K K K K	R T V L E S S S E E D D	E D D	
BAdV-4 Q997H7	P K K K K	R T V L E S S S E E D D	E D D	
OAV287 Q83907	P S K K K	K T I L E S S S E E E D	E D E D	
DAdV-1 O11422	P T K R K	R V V E I E S D E D E E	E D E D	
TAdV-1 NC_001958.1	P K K H S F K	Q A I V A Q H E V E S	D D D A F	C
FrADV-1 NP_062445.1	P K R H C Y K	S A V V C A Q E V E E	D E D V F	C

3.12. ábra. A DNS-kötő fehérjék C-terminálisának illesztése. Az egyes nemzetségeket vízszintes vonalak választják el, a savas aminosav-maradékok árnyékoltak, a NLS-okat és a konzervált fenilalaninokat keretezéssel emeltem ki.

ciszteinje helyén triptofán található,¹⁰¹ míg a másik motívum cink kötő aminosavai, csakúgy mint az OAV287-nél, konzerváltak.

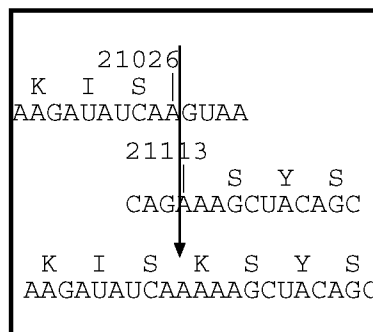
3.3.5. A késői fehérjék

A kapszid alkotóelemeinek összeállításához szükséges fehérjék

Ebbe a csoportba az 50/52K, a 100K és feltehetőleg a 33K fehérje tartozik, bár az utóbbinak a funkciója ismeretlen. A 33K fehérje mRNS-éről atadenovirális változatának esetében is megállapították, hogy splicing révén épül össze a kódoló szakasz¹⁸. A Rus 33K fehérje kódoló szakaszának vizsgálatakor azt találtuk, hogy splicing nélkül, az egyébként nagy biztonsággal megállapítható donor hely után minden leolvasási keretben stop kodonok következnek. A fehérje többi részével homológ peptidet kódoló szakaszt egy „splice” akceptor hely után megtaláltuk (3.13 ábra), tehát biztosra vehető, hogy a Rus-nál is történik splicing a 33K mRNS kódoló szakaszán.

Az 50/52K proteinek a HAdV-2 esetében 48 kDa tömegű prekurzorból képződő foszfoproteinek, melyek a virion összeállításában játszanak szerepet.¹⁰² Atadenovírusokban előforduló változataik mintegy 60 aminosavval rövidebbek, mint a mastadenovírusok esetében, és ez a Rus 50/52K homológja esetében is így van.

A 100K fehérje szerepe a II-es polipeptidek hexon fehérjévé való összeépülésének katalízisében van. A mastadenovírusoknál lévő változat mintegy 780 aminosavnyi átlagos hosszúságához képest az atadenovírusoké jóval rövidebb (DAdV-1 – 709 as, OAV287 – 629 as, Rus – 624 as). A homológ szakaszokon a többi nemzetség fehérjéivel mutatott hasonlóság 49–51%, ami nagy konzerváltságot jelent!



3.13. ábra. A 33K protein mRNS-ének kódoló szakaszán feltételezett splicing

A proteáz

A proteáz gén a kivétel nélkül minden adenovírusban megtalálható, mintegy kétszáz aminosavból álló, erősen megőrzött szerin-proteázt¹⁰³ kódolja. Méreténél és megőrzöttségénél fogva, és mert szekvenciája sok adenovírusból ismert, ideális filogenetikai számításokra.¹⁰⁴ Ezért parszimónia és távolság mátrix számítások felhasználásával törzsfákat szerkesztettünk a rendelkezésre álló szekvenciákból. Az eltérő módszerek hasonló eredményt adtak. Az ismert proteáz szekvenciákon végzett távolság-mátrix analízis eredménye a 3.3.5 ábrán látható. A törzsfa nincsen „gyökereztetve” (előre irányítva); a megjelenítéskor a a siadenovírusokat mi választottuk „outgroup”-nak. Az elágazásoknál feltüntettük a „bootstrap” értékeket. A függelékben nem szereplő fehérjék: BAdV-1, Peter Evans nem közölt eredménye; BAdV-6,¹⁰⁵.

Az ábrán látszik, hogy a javasolt *Atadenovirus* nemzetség tagjai (DAdV-1, PoAdV-1, BAdV-7, OAV287, BAdV-6, BAdV-4, Rus) jól elkülönülnek a többi mastadenovírustól. Bár a madár-, és a javasolt *Siadenovirus* nemzetségek csak két-két fajjal képviseltek, különállásuk szintén nem kétséges.

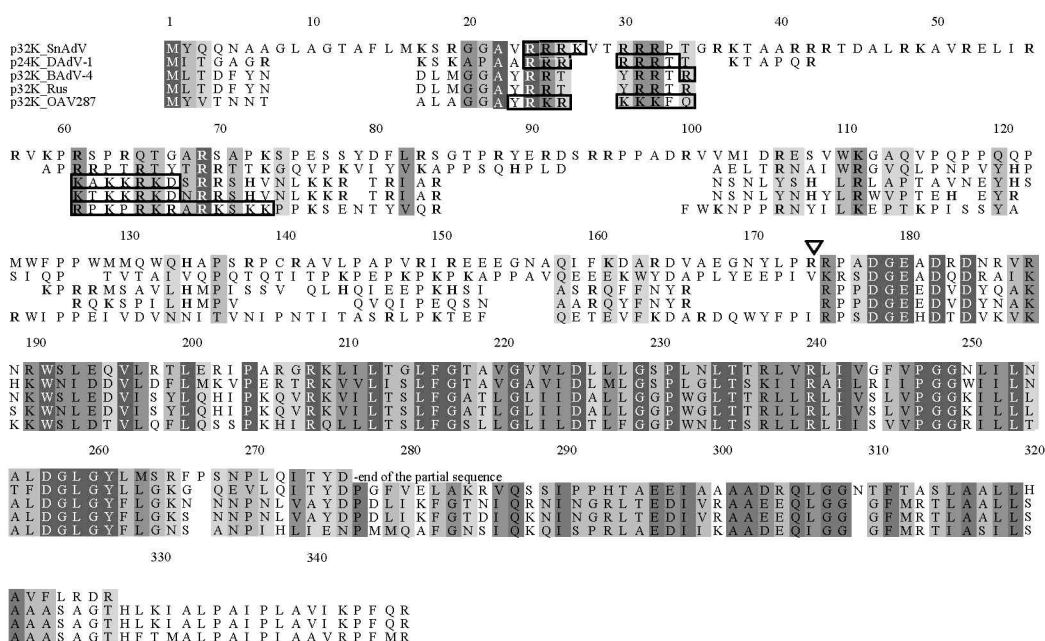


3.14. ábra. A proteáz szekvenciákban a távolság-mátrix módszerrel számított törzsfa.

3.3.6. Struktúrális fehérjék

A p32K protein

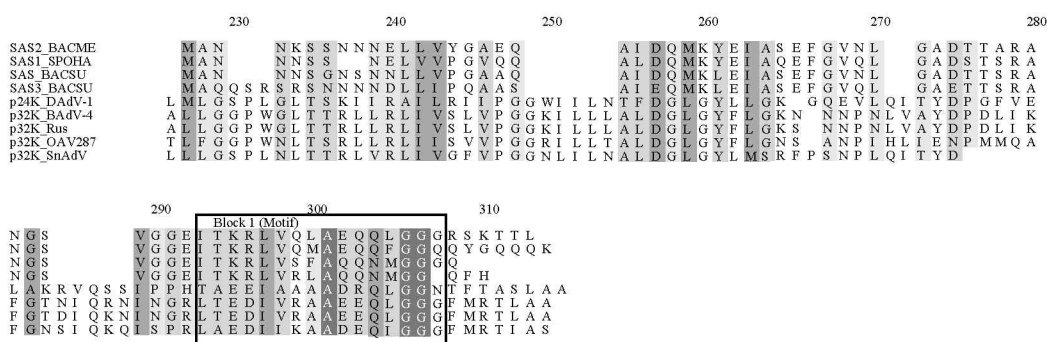
Ez a prekursor formában keletkező struktúrális fehérje az eddig vizsgált atadenovírusok mindegyikében így a Rus-ban is megtalálható; ugyanakkor más nemzetségek egyetlen tagjánál sem fordul elő. Rajta konzervált proteáz vágási helyek figyelhetők meg.¹⁸ Az ismert p32K homológ fehérjék illesztése a 3.15 ábrán látható.



3.15. ábra. A p32K fehérjék összehasonlítása. A két feltételezett domén határát üres háromszög jelzi, az N-terminális domén bázikus aminosavai vastag betűkkel jelöltek. A homológiák a megőrzöttségnek megfelelően árnyékoltak. A keretezett részletek a nukleáris lokalizációs szignálokat (NLS) mutatják. A szekvenciák elérési számai: p24K_DAdV-1: CAA70797, p32K_BAdV-4: NP_077384, p32K_Rus: AF238880, p32K_SnAdV: kígyó adenovírus részleges szekvencia, Farkas Szilvia nem közölt eredménye; p32K_OAV287: AAB19237

A rendelkezésre álló szekvenciákból készített illesztés alapján látszik, hogy a fehérje két eltérő szerkezetű részből áll. Az N-terminális vég hossza és szekvenciája erősen változó, benne sok bázikus aminosav található. A C-terminális rész ezzel szemben erősen konzervált. Mivel a p32K pontos funkciója nem ismert, és egyszerű Blast adatbázis kereséssel sem sikerült kimutatni

hasznoló proteinek, a szekvenciát a „Predict Protein”⁶⁸ web alapú alkalmazáscsomag segítségével elemeztettük. Ennek „MAXHOM”¹⁰⁶ nevű, távoli homológiák keresésére kifejlesztett programja a *Bacillus megaterium* egyik proteinjével való hasonlóságot jelzett, ami a a sporuláló Gram-pozitív baktériumok „kis savoldékony fehérjéi” (SASP)^{107, e} közé tartozik. Ezek az egymáshoz igen hasonló fehérjék a bakteriális kromoszóma spórába „csomagolásában”, és a környezeti hatásokkal szembeni védelmében játszanak szerepet. A MAXHOM által produkált eredmény alapján készíthető illesztés a 3.16 ábrán látható.



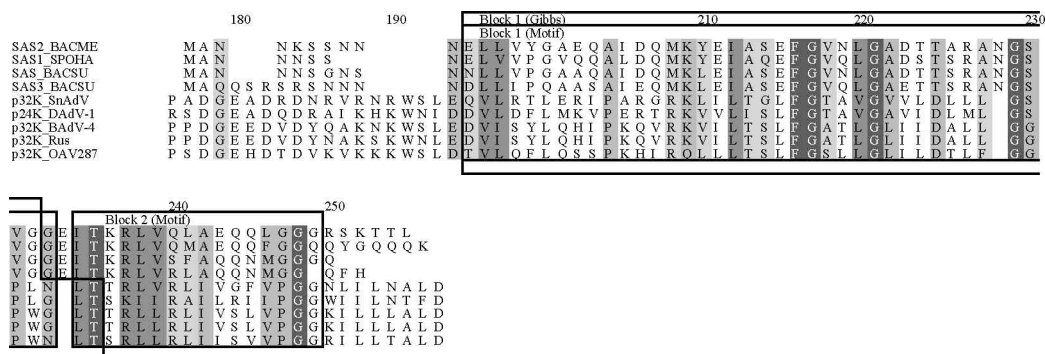
3.16. ábra. A p32K fehérjék és néhány SASP MAXHOM-eredmény alapján készített illesztése. A keretezett rész a „BLOCK MAKER” „Motif” alprogramjával talált blokkot jelöli. BACME: *Bacillus megaterium*, BACSU: *Bacillus subtilis*, SPOHA: *Sporosarcina halophila*. A szekvenciák génbanki elérési számai: SAS2_BACME: P10571,¹⁰⁸ SAS1_SPOHA: Q00213¹⁰⁹, SAS_BACSU: NP_390835,¹¹⁰ SAS3_BACSU: P02958¹¹¹

A Rus törzs p32K fehérje-szekvenciája mellett a többi homológ atadenovírus szekvenciát is felhasználva, a ClustalW⁶⁹ programmal sokkal jobb illesztést találtunk, ami a 3.17 ábrán látható.

Mivel a szekvenciák hasonlósága – a p32K esetében csak feltételezett – funkcióbeli rokonsággal látszott egybeesni, további hasonló szekvenciák felkutatásával megpróbáltunk magyarázatot találni a jelenségre. A két fehérjetípusból képezett, kilenc szekvenciából álló csoportban a „Block Maker”¹¹² alprogramjai által talált blokkok – amint a 3.17 ábrán látható – szinte teljesen átfednek. Ez arra utal, hogy a szekvenciák szignifikáns hasonlóságot mutatnak.

A kapott blokkokhoz hasonló részleteket tartalmazó szekvenciákat a „MAST” web-alapú motívum-kereső alkalmazással az NCBI Blast adatbázis-kereső

^esmall acid soluble protein



3.17. ábra. A p32K fehérjék és néhány SASP illesztése Clustal W programmal kapott eredmény alapján. A keretezett részek a „BLOCK MAKER” „Motif” és „Gibbs” alprogramjával talált blokkot jelölik.

programjai által is használt „nr” adatbázisban kerestük. Ehhez mindkét alprogram eredményeit egymástól függetlenül felhasználtuk. A nagyfokú hasonlóságot mutató fehérjék között mindkét esetben csak az adatbázisban található SASP és p32K homológok fordultak elő, de azok kivétel nélkül; míg a távoli homológiát mutató fehérjék a két keresésben különbözőek voltak, vagyis bár a blokkok alkalmasak a hasonló szekvenciárészletek kimutatására, más csoportba tartozó, számottevő mértékben hasonló fehérjéket nem találunk. OAV287 és a BAdV-4 p32K fehérjékben viszont a program különös módon két „Motif 2.” blokkot is azonosított, melyek közül a nehezebben felismerhető, amint az a 3.16 ábrán látható is, a MAXHOM által „homológoknak” vélt területre esik.

Miután a SAS proteineken és a p32K homológokon kívül más hasonló szekvenciájú fehérjét nem találtunk, így semmit nem tudunk mondani arról, hogy a két fehérjecsoport hasonló részei között van-e valamilyen filogenetikai kapcsolat. Meglehet, hogy függetlenül fejlődtek hasonlóvá hasonló funkciók betöltésére, vagyis az adott „szervezet” DNS-ének valamilyen szűk kompartmentbe (vírus kapszid, bakteriális spóra) való „csomagolására”, és annak a környezet DNS-károsító hatásaival szembeni védelmére. Lehetséges azonban, hogy a p32K őse egy olyan bakteriális fehérje, mely az adenovírusok bakteriofág őseinek közvetítésével¹¹³ az egyik adenovírus nemzetségben máig fennmaradt. Az OAV287, Rus és BAdV-4 p32K szekvenciájában talált két-két Motif 2. blokk jelenlétét legkönnyebben az adott génszakasznak a távoli múltban bekövetkezett duplikációjával lehet magyarázni. A genomvéghez közeli szekvencia-duplikációk nyomát így a bal végen is felfedezhetjük.

Az irodalomból számos leírás ismert diagnosztikai célokra kifejlesztett PCR-primerekről, melyek különböző vírusok, köztük adenovírusok kimuta-

tására is alkalmasak.^{114;115;116} A C-terminális szakasz konzerváltságát látván úgy döntöttünk, hogy megkísérelünk a nemzetségre specifikus gén e darabjára olyan PCR-primereket tervezni, amelyekkel e vírus-nemzetség minden tagjának jelenléte kimutatható.

Célunk volt emellett az is, hogy a kapott DNS-darabok szekvenciáit – amennyiben lehetséges – faj-azonosításra, illetve az egyébként ismeretlen vírusok rokonsági viszonyainak előrejelzésére is felhasználjuk. (Komolyabb filogenetikai vizsgálatok elvégzésére ez a felerősített szakasz rövidsége miatt természetesen nem alkalmas.)

A tervezett degenerált primerekkel (p32for: 5'-GAYGGAGAASMIGAYVIRGA-3' p32rev: 5'-CCTCCKAKYTGTYSYDMTKC-3') hét különböző kérődző eredetű adenovírus DNS-éből erősítettünk fel fragmentumokat. Ezek között szerepelt két, részletesen még nem jellemzett adenovírus izolátum is. Az „Ibex” törzset egy Ibex-házi kecske hibridből izolálták,¹¹⁷ a 24BH elnevezésűt pedig korábban a laboratóriumunkban egy primer borjú here szövettényezetben kontaminációként találtuk^f. A kapott DNS fragmentumok illesztési képe a 3.18 ábrán látható. Az itt szereplő vírusokon kívül az DAdV-1 DNS-t tartalmazó mintából is fölszaporítottuk a homológ DNS-darabot.

Az eredmény alapján kijelenthetjük, hogy a primer-pár, amit a C-terminális szakaszok sokszorozására terveztünk – legalábbis az általunk vizsgált esetekben – alkalmas volt a vizsgálandó génszakasz fölerősítésére. Az amplifikált szakasz alapján készített, a 3.19 ábrán látható távolság alapú törzsfából megállapíthatjuk, hogy az erős konzerváltság ellenére a vizsgált szakasz, bár a fajok azonosítására nem alkalmas, a vírusok nagyobb csoportokba való sorolására megfelel. Ezáltal lehetőség nyílik arra, hogy egy egyszerű és rövid vizsgálat alapján képet alkossunk a vizsgálandó vírus nemzetségen belüli helyéről. A fenti kísérletben például három csoportot sikerült biztonsággal elkülöníteni: a Rus törzset, a BAdV-7-et, 24BH-t, és BAdV-4-et magában foglaló csoportot, valamint a BAdV-5-öt, BAdV-8-at, és az „Ibex” izolátumot tartalmazót. Így a Rus izolátumról megállapíthattuk, hogy megbízhatóan elkülönül a BAdV-7 – 24BH csoporttól, noha eredetileg BAdV-7 szerotípus egy törzseként írták le.

A hexon

Az atadenovírusokra jellemző a II-es protein génjének elhelyezkedése is az öt körülvevő génekhez képest. A pVI és a II kódoló része közti távolság csak néhány nukleotid, míg a proteáz-gén kódoló részének vége átfed a hexon

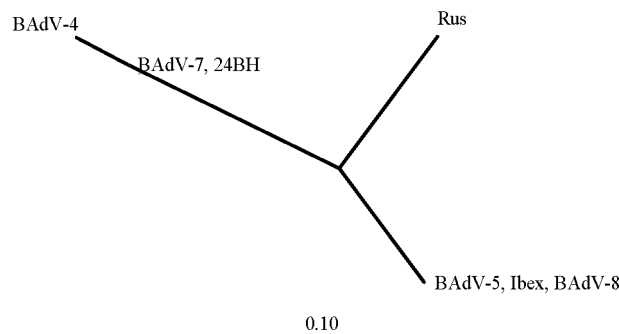
^fBenkő Mária nem közölt eredménye.

	10	20	30	40	50
Rus	ATAATGCTAAAGCAAATGGAACCTTGAAGATGTCATTTCCTATTACAA				
BAdV--4	.C.A.C...A...G...A...A...G				
24--BH	--CAA.C...A...G...A...A...G				
BAdV--7	-----A...A...G				
BAdV--5	C.....G				
BAdV--8	C.....G				
Ibex	C.....G				
Rus	CACATTCCGAAGCAAGTCAGAAAAGTGATTTTAACAAGCTTATTTGGCGC				
BAdV--4	T.....A...AC.....				
24--BH	T.....A...AC.....				
BAdV--7	T.....A...AC.....				
BAdV--5	.G...T.....A...G				
BAdV--8	.G...T.....A...G				
Ibex	.G...T.....A...G				
Rus	CACCTCTAGGCTCTAATAATAGATGCTTTACTGGGAGGTCGTTGGGGTTAA				
BAdV--4	...C.T.....T.....				
24--BH	...C.T.....T.....				
BAdV--7	...C.T.....T.....				
BAdV--5	...T.G.....C...C...G				
BAdV--8	...G.....C...C...G				
Ibex	...G.....C...C...G				
Rus	CAACACGACTGCTTCGCCTTATAGTTTCTTTGGTTCTCGGCGTAAAAATT				
BAdV--4	T...G.....A.....				
24--BH	T...G.....A.....				
BAdV--7	G...T...G.....A.....				
BAdV--5	...G..C...G..G.....G.C.C.A				
BAdV--8	...G..C...G..G.....G.C.C.A				
Ibex	...G..C...G..G.....G.C.C.A				
Rus	TTACTCTAGCTTTGGATGGCTTAGGGTACTTTCTCGGAAAAAGCAACAA				
BAdV--4	...A.....A.C..A.....T.....A.....				
24--BH	...A.....A.C..A.....T.G.....A.....				
BAdV--7	...A.....A.C..A.....T.G.....A.....				
BAdV--5	...A.....A.T..C.....T.....C.G..C.....T.....				
BAdV--8	...A.....A.T..C.....T.AC...C.G..C.....T.....				
Ibex	...A.....A.T..C.....T.AC...C.G..C.....T.....				
Rus	TCCAAATCTTGTTCATATGATCCAGATTTAATAAAGTTTGGAACTGATA				
BAdV--4	CC.G..T...C...A..C				
24--BH	G..CC.G..T...C...A..C				
BAdV--7	G..CC.G..T...C...A..C				
BAdV--5	T..A..G.....				
BAdV--8	T..A..G.....				
Ibex	T..A..G.....				
Rus	TACAGAAAAACATTAATGGTCGCCTAACGGAAGACATAGTGCST				
BAdV--4	.C..A..G.....T.....AGC..A				
24--BH	.C..A..G.....T.....A...				
BAdV--7	.C..A..G.....T.....A...				
BAdV--5	.C.....C..T.....A..G.				
BAdV--8	.C.....C..T.....A..G.				
Ibex	.C.....C..T.....A..G.				

3.18. ábra. A p32K gén PCR segítségével felerősített darabjainak illesztése: Ibex: ibex-házikecske hibridből izolált, nem szerotipizált vírus, 24BH: szövet-tenyészetből izolált ismeretlen vírus, „-”: hiányzó szekvenciariész, „.”: egyezés a Rus szekvenciával

kódoló szakaszának elejével.¹⁰⁴ A Rus törzsnél is hasonló az elrendezés, bár az átfedés mindössze egy nukleotidnyi (3.1 ábra).

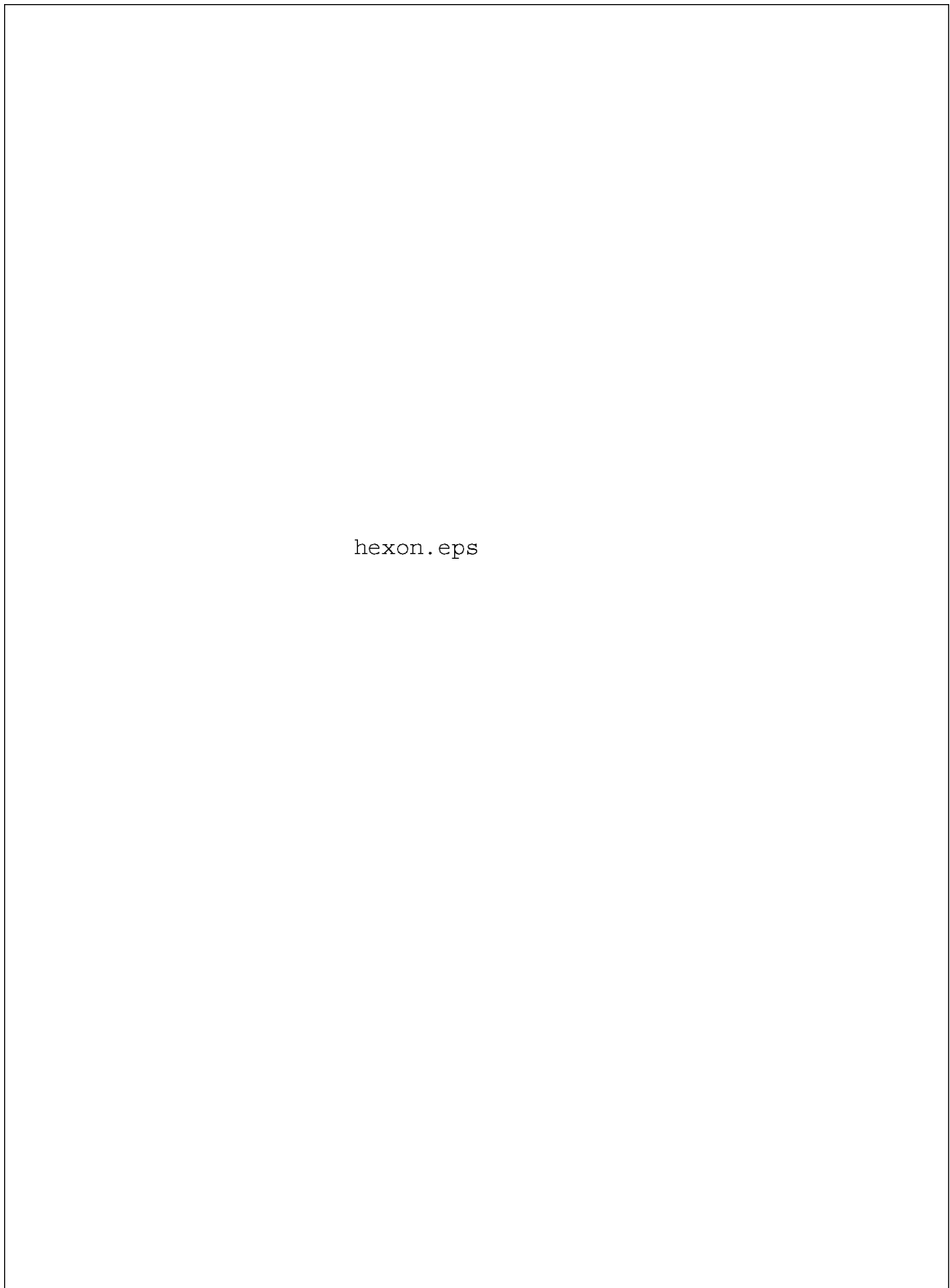
A hexon protein az adenovírusok egyik legjellemzőbb fehérjéje. Három azonos polipeptidből áll, és mint a vírus-kapszid fő alkotóeleme, igen erős megőrzöttséget mutat a víruscsaládban. Génjének szekvenciája sok adenovírus törzsből ismert. Röntgenkrisztallográfiás úton a HAdV-2 hexonjának három dimenziós felépítését is meghatározták.¹¹⁸ Eszerint a nagy sűrűségű alapi részt főként β -redőkből álló hengerek, a csúcsi részt pedig három hurok építi fel. Ez utóbbiak a kapszid külső felszínén találhatóak. A elsődleges szekvencián az alapi résznek megfelelő két erősen megőrzött egységet a hurkok variábilis és hipervariábilis szakaszai szakítják meg^{119;120}. Mivel a hexon gén



3.19. ábra. A p32K gének amplifikált C-terminális részéből Juke-Cantor távolságindex alapján, Fitch–Margoliash algoritmussal számított törzsfá. A 24–BH és *Ibex*¹¹⁷ nem jellemzett izolátumok

szekvenciája igen sok adenovírus típusból ismert, lehetőség van a nemzetiségekre jellemző sajátosságok kiemelésére.¹²¹ A fenti jellemzőket figyelembe véve megállapítható, hogy a hexon polipeptid kiválóan alkalmas filogenetikai számításokra. A hexon aminosav-szekvenciákon végzett távolság-mátrix számítás eredménye a 3.3.6 ábrán látható. Az ágak hossza arányos a törzsek közötti¹⁰⁴ távolsággal. A „bootstrap” értékeket – az emberi adenovírusok kivételével – feltüntettük. A fát „gyökerezítettük”: a siadenovírusokat választottuk „outgroup”-nak, mivel ezek tekinthetők a legősibb adenovírusoknak a kételtűekből való származás alapján⁵⁴.

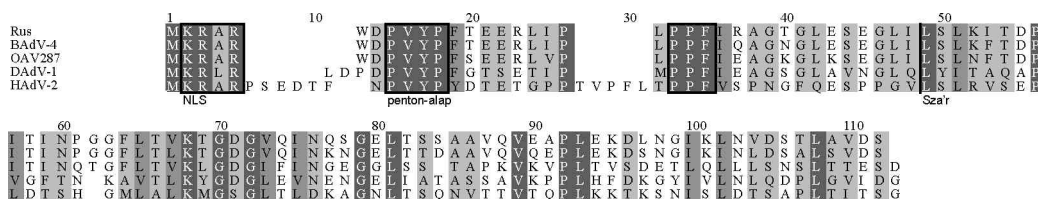
Meg kell jegyezni azonban, hogy a távolság-mátrix számítások csak az egyes vírusok közti távolság mértékét határozzák meg valamilyen paraméter-táblázatot felhasználva. A számítások alapján készített távolság alapú törzsfák ily módon a leszármazásra vonatkozólag nem nyújtanak biztos információt. Mivel azonban az ághosszak arányosak a vizsgált szekvenciák közti távolságokkal, az ilyen fák nagyon alkalmasak a különböző csoportok közti távolságok szemléltetésére. Ezért választottuk ezt a számítási és ábrázolási módot. A 3.3.6 ábrán bemutatott törzsfán is látszik a proteáz enzim alapján számított fához hasonlóan (3.3.5 ábra), hogy a vírusok az egyes nemzetségeknek megfelelően oszlanak meg. Leszűrhetjük tehát azt a tanulságot, hogy a genomszerkezeti különbségek alapján készíthető felosztást alá lehet támasztani a konzervált szekvenciákon végzett filogenetikai számításokkal is.



3.20. ábra. A hexon fehérjéből távolság-mátrix módszerrel számított törzsfa.

A fiber

A Rus vírus fiber fehérjéje viszonylag rövid. A DAdV-1-hez és az OAV287-hez képest meglehetősen kis méretű a feji része. A farki szakaszon a szekvencia igen erősen konzervált, ahogy az a 3.21 ábrán is látható. A magi lokalizációért felelős szignálon (NLS) és a penton-alaphoz kötődő részen kívül még egy különös, erősen konzervált szakasz található itt, aminek a funkciója ismeretlen.



3.21. ábra. Fiber fehérjék farki részének illesztése

A szár rész nagyjából 15 aminosavnyi ismétlődéseit^g a legújabb modellnek²² megfelelően a 3.22 ábrán tüntettük fel. A HAdV-2-nél leírthoz hasonlóan a harmadik ismétlődés „szabálytalan”, de ezen túl a Rus-nál, ebben a szerkezetben a „szoros fordulóra” hét esetben nem prolin, vagy glicin esik. Ez azt jelenti, hogy ha a szár a modellben leírtaknak megfelelően épül fel, az ismétlődések harmadánál nagy valószínűséggel akkor is más szerkezet alakul ki. Megemlítendő még, hogy a hidrofób mag aminosavainak illesztéséhez az ismétlődések második pozíciójába néhol hézagokat kell illeszteni.

A pIIIa és a penton-alap

A pIIIa protein feltehetőleg a peripentonális hexonok és a VII-es fehérje közti híd szerepét tölti be. Az emberi adenovírusoknál jellemző rá az „(MLI)xG†G” típusú proteáz vágási hely,¹²² helyett az atadenovírusoknál a másik típust (M/L/IxGx†G) találjuk a C-terminális közelében. A Rusnál kettő fordul elő, ezek közül a 490. pozíciónál található „LFGD†G” a nemzetiség eddig ismert többi tagjában is megőrzött.

A penton alap (III-as protein) a HAdV-2-nél végzett vizsgálatok alapján fontos szerepet játszik a vírus sejtbe jutásában. A HAdV-2-ben, és még néhány másik mastadenovírusban találtak egy „RGD”^h motívumot, ami a sejt integrinjeihez történő kapcsolódáshoz fontosnak tűnt¹²³. Kiderült azonban,

^g„pseudo-repeat”

^harginin-glicin-aszparaginsav

<i>Farok</i>	
MKRARWDPVYPFTEERLIPLPPFIRAGTGLESEG	
<i>Szár</i>	
	1 2 3 4 5
1	LILSLKIT DPITINPG
2	GFLTIVKTG DGVQINQS
3	GELTSSAAVQVEAPLEKDLN
4	G IKLNVDS TLAVDSA
5	GKLTVALN NPLENTLT
6	G INLKLSN SFSIDDA
7	GKLDIHSPE LPLSITPE
8	GKLFQLS NCFASDVN
9	GKLVLPPE LPLSLTPE
10	GKIFQLS NSFSTDVN
11	GKLSLAFPK APLSVNQE
12	GKLFQLS NSFNIDS
23	GKLVLTARNF HFLSVLKKAS
14	FKIRTSFNGR WPFNFKNRKSFGSKR
15	RSLKIKLS NPFEEELN
16	GNLTIKTN YPTYVDDV
17	GFLNFKVG EPFTVQN
18	DNLALSLNNS LSIDN
19	GKLAVKTA GPIQSTDQ
20	G IKLNVA NPFSTSSNL
21	LSLKIT SPLTINTS
22	GELTTITR QGSCIVSFMD
23	FIIALGWQV IPSSQRF
24	IYILNCSQF MPTSNVNSI
<i>Fej</i>	
YFQADNGLENVFIVDTPFYTSVTQQMPDKSIRTYG	
VSISKNTTISINFSSLEPNVMVSSWTASITRL	

3.22. ábra. A Rus fiberének aminosav-szekvenciája. A szárrész ismétlődéseinek számozott pozíciói: hidrofób mag: 1, 2; széli hidrofób részek: 3, 5; „szoros forduló”⁴.

hogy ez a motívum még a mastadenovírusok között sem fordul elő általánosan,^{124;19;125;21;20;126;127} és a Rus esetében is hiányzik.

A pVI, pVII, pVIII és pX fehérjék

Ezek a virion kapszidjának összeállításában és rögzítésében fontos szerepet játszó kis szerkezeti fehérjék prekuzorként keletkeznek, és a virális proteáz hatására érnek funkcionális proteinekké.

A pVI fehérje fontos struktúrális része a virionnak, a vírusmagot rögzíti a kapszidhoz. A Rus pVI proteinen is a többi homológ proteinre jellemző két proteáz vágási hely található: a 27. pozíciónál bekövetkező vágáshoz az „LRG†G” hely, a C-terminális közelében pedig másik típusú „MTGD†G” hely. Az utóbbi vágás eredménye a „GVNSVTRRYCY” peptid (VIc), ami a virális proteáz aktiválásában játszik kulcsszerepet¹²⁸. Szekvenciája megfelel a javasolt „GVxxxxxxCY” konszenzusnak²⁰.

A pVII a DNS-sel képez komplexet, felépítése a többi adenovíruséval egye-

zik meg: a konzervált N-terminálist egy pozitív töltésű aminosavakban gazdag, de egyébként nem megőrzött C-terminális követi, ami viszont az atadenovírusokra jellemzően jóval rövidebb a mastadenovírusokénál. A pVII-nek az Atadenovírus nemzetségre jellemző sajátosságának tekinthetjük azt is, hogy az ezekre a fehérjékre jellemző két konszenzusos proteáz vágási hely közül az egyik minden ilyen vírusból ismert proteinben megváltozott¹⁸: a (M/L/I)xG†G típusú vágási szekvencia helyett (M/L/I/)xAx†G szekvenciát találunk, ami az (M/L/I/)xGx†G típusú hely egy változatának tekinthető (3.23 ábra).

A pVIII protein a hexon-alegységek belső szilárdításában játszik szerepet. A „vírusmagot” köti össze a kapsziddal. N- és C-terminális része megőrzött, rajta a Rus törzs esetében két potenciális proteáz vágási hely is található, és ezek a nemzetség többi tagjában is megvannak: a 117. pozíciónál (LTGT†G) és a 147-nél (LQGM†G). A mastadenovírusoknál csak az utóbbi megőrzött.

A pX az adenovírusok kevésbé konzervált, kisméretű és esszenciális komponense, a kapszidon belüli pontos helye nem ismert. Belőle képződik a μ peptid is, ami a DNS-hez közvetlenül kapcsolódik. A látszólagos ellentmondás abból adódik, hogy nagy részét bázikus aminosavak halmaza alkotja, konkrét aminosav szintű megőrzöttség csak a C-terminális közelében van. Az atadenovírusoknál – a mastadenovírusokkal szemben¹²⁹ – csak egy proteáz felismerő helyet találunk a második típusúból, jelen esetben a 38. pozíciónál található vágással (LTG†G).

```

1      10      20      30      40      50
HAdV-2 MSILILISPSNNTGWGLRFPSSKMFGGGAKKRS DQHPVVRVGRGHYRAAPWGAHKRGRTGR
BAdV-3 MAAILISPSNNTGWGLGTCN.KMYGGGARIRSDLHPVKVRSHTYRAAWGSR.TGRVGRRA
CAAdV-1 MAAILISPSNNTGWGLGTH.KLFGGGAKKRS DQHPVVVOAHYRAPWGSKGR.RRRPGR.A
MAdV-1 MSILILISPSDNTGWGLGTTG.KMYGGGARKKRS AEHPVHVRSYWRAPWGS.RNRRRVATVA
BAdV-4 MSILILMSPSDNTGWGSIGTALMRATGVKFSKKRQPVVRVPYYRAQWQQLNARTPLAKL
Rus MSILILMSPSDNTGWGNIGTSSMRATGLKFSKKRQPVVRVPYYRAQWQQLNARTPLAKL
OAV287 MSILILVSPSDNTGWGLGTRMLRATGLRFSEREPEVRVRSYWRAPWQQLNGRKSAARQL
DAdV-1 MSILILMSPADNTGWGLGTRMLRATGLRFSEREPEVRVRSYWRAPWQQLNGRKSAARQL

ML/LNxxG ML/LNxxG

60      70      80      90      100      110      120
VDDAI DAVVEEARNYTPTPPPVS TVDAAIQT VVRGAR RRYAKMKRRRRRVARRHRRRP GTAAQRAAAA
TAALADAVAATG. . . . . DPVADTIEAVVADARQYRRRRRRGVRRVRRLRRSP RTALQRRVR
RGVPLDPKTEAE. . . . . VVATIDEVARNGPPAARLVLEAARRVGAYNLR. . . ARKLT
AEDAEAPQLEDVAQAP. . . . . ATVPIVRRHRRRVGGSARTRGLRKSAVRRAAIVRAVNGAAAAAA
RENLLKKYERKHKRKR. . . . . QKKKTQTVFV. . . . . EDSPPP RRRR.RRENK
RENLLKKYERKHKRKR. . . . . QKKKTQTVFV. . . . . EESPPP RRRR.KEEIK
KTKLKYYEKLYRDR. . . . . LKR.KTVVP. . . . . KKKRSP. . . . . TSP
RRGYKRYQAI HKGRQL. . . . . AKQRRRAVL SRANRHAAATNHRRQLQLVAE RRAAI AESRRRRQA

130      140      150      160      170      180
ALLNRARRRTGRRRAAMRAARRLAAGI VTVP PRS RRRRA AAAAAAAI SAMTQGRRGNVY. . . WVRDS
SVR. . . RQVAR. . . ARRVGRRAAA. . . . . I AADAAAMAMAAPAR. RRRNIY. . . WVRDA
P. . . . . AGRAMAAMRA. . . . . RQM
PTVPASANFAAMVGAI ANARAAAYRHRRS RAGI NPVP ASRS TTRVRLPRTVRFHPSMGAFHRRVWRI
TRR. . . . . IEKYLDALAK. . . . . I RA
HDV. . . . . LKKYLDALAK. . . . . I KA
ADR. . . . . LKKYLDALAK. . . . . I KA
AAA. . . . . AQEVVNAAVEAAVEAAAAAQPAADEAP KKK

190      200
VSGLRVPVVRTRPPRN. . .
ATGARVPVTTTRPTVSN TV
VNQAQR. . . RKRRRVRSK. . .
HTRGRRKASVRRRRRT. . .
FNRRAR. . . RRAAI. . .
FNRRAR. . . RRAAI. . .
FNRRAR. . . RRAAI. . .
SNRKRGGPPKRLNI. . .

```

3.23. ábra. Néhány pVII protein illesztése A proteáz vágási helyeket keret-
zéssel emeltem ki. A fehérjék: HAdV-2 (P03266), BAdV-3 (NP_046319),
CAAdV-1 (P35989), MAdV-1 (NP_015542.1), BAdV-4 (NP_044707),
OAV287 (AAA84976), DAdV-1 (NP_044707)

4. fejezet

Összefoglalás

Munkánk során meghatároztuk a Rus szarvasmarha adenovírus izolátum szinte teljes genomjának bázissorrendjét. A kapott szekvenciák számítógépes analízisével megállapítottuk a feltételezett fehérjék aminosav-szekvenciáit. A genomszerkezet és a kódolt fehérjék összehasonlító elemzésével bizonyítottuk, hogy a Rus törzs általános genomszerveződése eltér a *Mastadenovirus* nemzetségnél megszokottól, viszont nagyon hasonló a DAdV-1, OAV287 és BAdV-4 vírusoknál tapasztaltnak. Erre a vírusra is jellemző a többi adenovírus csoporttal homológ gének viszonylagos rövidsége, a kódoló szakaszok közötti szokatlanul kis távolság, az E1B és E4 régiók eltérő összetétele, valamint az E1A, V, IX gének és a mastadenovírus-szerű E3 régió hiánya.

Az ITR szekvenciák összehasonlításából megállapítottuk, hogy azok összetétele és hossza a javasolt nemzetségekre jellemző eltéréseket mutat, amiből arra következtethetünk, hogy az adenovírusok DNS-replikációjáról főként a HAdV-2 és HAdV-5 vizsgálata alapján szerzett ismeretek egy jelentős része nem általánosítható.

Az LH korai régiót elemezve megállapítottuk, hogy az polimerázzal átfedően kódolt, korábban különböző adenovírusoknál leírt feltételezett fehérje valójában „műtermék”, az LH2 fehérje az összes atadenovirális megfelelőit figyelembe véve nem áll leszármazási kapcsolatban a mastadenovirális E1B 19K fehérjével, az LH3 pedig távoli, de határozott rokonságot mutat az E1B 55K mastadenovírus fehérjével.

Megtaláltuk a szintén csak erre a nemzetségre jellemző ún. p32K protein génjét. A szekvenciahomológiák keresése során azonosítottunk egy bakteriális protein-családot, amely a p32K feltételezett funkciójához hasonló szerepet tölt be a bakteriális sporuláció során, és emellett bizonyos alacsony fokú hasonlóságot mutat szekvencia-szinten is.

E fehérje C-terminális konzervált részét kódoló DNS-szakaszra olyan, di-

agnosztikai célra alkalmasnak látszó PCR-módszert dolgoztunk ki, amelyekkel a birtokunkban lévő kérődző eredetű adenovírus mintákból sikerült megsokszoroznunk a kívánt DNS-darabot. A PCR termékek szekvenciája alapján az atadenovírusok közé soroltunk két, eddig molekuláris módszerrel még nem jellemzett adenovírus izolátumot is. Megmutattuk, hogy a kapott szekvenciák alkalmasak a vizsgált atadenovírusok csoportokba rendezésére, noha ez a csoportosítás nem szükségképpen esik egybe más géneken végzett filogenetikai célú vizsgálatok eredményeivel.

A hexon és a DNS-polimeráz fehérjéken végzett filogenetikai számításokkal megmutattuk, hogy az *Adenoviridae* családba tartozó vírusok négy, egymástól viszonylag távol álló csoportot alkotnak. Az a tény, hogy különböző fehérjék alapján végzett ilyen irányú vizsgálatok eredményei a nagyobb csoportok szintjén gyakorlatilag megegyeznek a genomszerkezeti eltérések alapján végezhető csoportosítással, alátámasztja a víruscsalád újrafelosztásának szükségességét¹³⁰.

Ugyanezekből a vizsgálatokból, valamint a p32K protein gén homológok C-terminális végeinek DNS szekvenciáin végzett filogenetikai számítások alapján megállapítottuk azt is, hogy az általunk vizsgált Rus törzs nem sorolható a BAdV-7 típusba, mint azt korábban leírták,⁵⁵ hanem egy új vírus-típus. Eddigi vizsgálataink alapján azt szűrhetjük le, hogy az ismertté vált atadenovírusok közül a BAdV-4 vírussal mutatja a legközelebbi rokonságot.

Az úgynevezett E3 régió már ismert szekvenciájának analízise során meghatároztuk a rajta potenciálisan kódolt három fehérje aminosav-sorrendjét. Az így kapott aminosav-szekvenciákból megállapítottuk, hogy a vizsgált proteinek egy, az atadenovírusokra általánosan jellemző, fehérje-csoport képviselői. A nukleotid- és aminosav-szekvenciák elemzésével megmutattuk, hogy ezek a fehérjék a genomok jobb végén lezajlott duplikációk során alakultak ki. Az elemzések alapján felfedtük a Rus vírus e régióján tapasztalt duplikációk valószínű sorrendjét is.

5. fejezet

Summary

The adenoviruses comprise one of the most important virus family of vertebrates, especially from scientific point of view. They were first reported as distinct pathogenic agents in 1953. The finding that certain human adenoviruses can induce tumors in newborn hamsters made them the model of tumorigenesis instantly. Many of the steps of their replication have been studied as models of certain intracellular processes. The research of their transcription regulation revealed the maturation processes of the eucaryotic mRNA. Their DNA replication, involving relatively few factors made them ideal for the study of DNA replication in general. The protein primed DNA-synthesis was first studied in adenovirus systems. Numerous research techniques first applied to map the genomic localization of adenoviral mRNA have been applied subsequently to many other pro- and eucaryotic systems. Transport of proteins from the cytoplasm to the nucleus, assembly of virions and of nucleic acid-protein complexes were studied in adenoviruses too. Nowadays, the research of the interaction of adenoviral proteins with each-other and with different transport mechanisms and life cycle regulator proteins of the host cells are important topics of the studies on adenoviruses. Their relatively harmless nature, easy replication and well understood replication processes made them important candidate vectors also in gene therapy research. Due to these promising characteristics, several differently designed generations of adenovirus based vectors were developed in the past decades.

However, although the scientific literature dealing with adenoviruses could fill libraries, most of the results were obtained from a few human adenoviruses, mostly from HAdV-2 and HAdV-5. Concerning animal adenoviruses, in most of the cases the researchers only serotyped the isolated strains, or in case of new host organisms electron micrographs were often made without further molecular studies, assuming that the virion similar to HAdVs covers

a similar genome, with the same overall characteristics. If we look up the results of the sequencing projects, we can see, that the first complete genome sequence of a human adenovirus (HAdV-2) was determined in 1986, and the genome sequences of the first animal adenoviruses were published a decade later.

With the spread of automated DNA sequencers, the situation changed dramatically in the last 6 years. Several complete genome sequences of various animal adenoviruses are determined year after year, and the amount of partial sequence data is increasing rapidly too.

However, despite the fact that the occurrence of adenoviruses was observed in each vertebrate class, the official taxonomy of this virus family still recognises only two genera, (*Aviadenovirus* and *Mastadenovirus*).

A proposal for the new classification of *Adenoviridae* was made recently based on the different characteristics of the four distinct virus groups. The new classification of the family *Adenoviridae*, according to the proposal, divides the adenoviruses to four genera. Besides the two conventional genera the establishment of two new genera was proposed, *Atadenovirus* and *Siadenovirus*. Members of genus *Atadenovirus* lack the E1A, V, and IX genes, and the E3 region completely. These viruses have a characteristic structural protein named p32K near to the left end of the genome, a putative E3 analogue region on the right end of the genome and an altered E4 like section between the fiber gene and the „E3” region. Candidate atadenoviruses are the subgroup 2 bovine adenoviruses, the duck adenovirus 1 (DAdV-1) and the ovine adenovirus isolate OAV287. The members of the genus *Siadenovirus* are the turkey adenovirus 3 (TAdV-3) and the frog adenovirus 1 (FrAdV-1), which surprisingly share a lot of important characteristics: a sialidase-like gene near to the left hand end of the very short (26 kilobasepairs) genome, lack of E1A, and E3 homologues and an E4 region with one gene.

The genomic organization of Rus proved to be identical to that of OAV287 and BAdV-4, and very similar to DAdV-1. The base composition of the 28482 bp long sequence was also in the range of atadenoviruses (64.4% AT). Comparing the inverted terminal repeat (ITR) sequences of all the examined adenoviruses, we have concluded that the length and organization of the ITRs are characteristically different in the four proposed genera. Thus, the majority of the knowledge on the mode and mechanism of adenoviral DNA replication compiled mainly by studying HAdV types 2 and 5, might not necessarily be found universally true for animal adenoviruses, especially for those from birds or lower vertebrates.

The analysis of the early region LH (on the left hand end of the Rus genome) revealed that the putative protein the gene of which overlaps the

DNA polymerase gene (encoded by the complementary strand) is indeed an artifact. Furthermore, the protein product of the LH2 gene is not phylogenetically related to the 19K protein of the mastadenoviral E1B, while LH3 exhibits a weak but sound homology with the E1B 55K protein.

We have been able to identify the gene of p32K, a precursor protein uniquely found in the genome of all atadenoviruses studied so far. Attempts to find homologous proteins in the databases by conserved motif searches resulted in the recognition of bacterial protein group, i.e. the small acid soluble proteins (SASPs) that is involved in packaging the DNA into the bacterial spores. Since atadenoviruses do lack an important mastadenoviral core protein (V), hypothesizing functional relatedness between p32K and SAS proteins seems to be reasonable.

Degenerate oligonucleotide primers, designed on the basis of the amino acid alignment of all known p32K sequences, were synthesized and tested in PCR. The method seemed to be feasible for the genus specific detection of atadenoviruses.

Distance matrix analyses performed with the hexon and protease gene sequences showed clustering of members of the family Adenoviridae into four, clearly separated clusters. This separation, which corresponded to the four different types of genomic organization found in the members of the two conventional (Mast- and Aviadenovirus) and two proposed (At- and Siadenovirus) genera, further supported the need of reclassification of the family. Similar analysis on the p32K sequences of different bovine atadenoviral serotypes revealed that strain Rus is not an isolate of BAdV-7 (as thought and published earlier), but represents a new type. From the data at hand, we have concluded that BAdV-4 seems to be the closest relative of Rus.

We have predicted the amino acid sequence of three proteins encoded by the putative "E3" region of Rus. Analysis of the amino acid sequences showed that these proteins are the representatives of a protein group generally characteristic for atadenoviruses. By nucleotide analysis, we have demonstrated that these genes might have been evolved by duplications or reiterations at the right hand end of the genome. The likely temporal sequence of these duplications was also established.

6. fejezet

Köszönetnyilvánítás

Szeretnék köszönetet mondani témavezetőimnek Benkő Máriának és Harach Balázsnak, a kapott témáért, a szakmai irányításért, és azért, hogy munkámat a laboratóriumunkra oly jellemző, barátságos légkörben lehetővé tették.

Köszönöm munkatársaimnak Dán Ádámnak, Farkas Szilviának, Kovács Gábornak és Ursu Krisztinának a közös munkát, a dolgozat javításában és bírálataiban nyújtott segítségüket. Külön köszönetet érdemelnek a dolgozat azon részeinek technikai kivitelezésében nyújtott támogatásukért, amelyekhez Microsoft termékek igénybevételére volt szükség. Dán Ádámot külön köszönet illeti, hogy a BAdV-4 genomjának szekvenálásánál szerzett adatokkal, és a róla szóló Ph.D. dolgozatával is segítséget nyújtott a Rus törzs szekvenációjának feldolgozásához. Őt és Ursu Krisztinát elismerés illeti azért is, hogy az Országos Állategészségügyi Intézettel kooperációban a DNS-szekvenálást lehetővé tették és benne sokat segítettek.

Kovács Gábort a filogenetikai számítások elvégzésében nyújtott fontos segítségét szeretném külön is kiemelni.

Köszönet illeti Egyed Lászlót tanácsaiért és hasznos észrevételeiért.

Végül szeretném megköszönni feleségemnek Vibók Ildinek, hogy szorgos munkával rengeteg hibát gyomlált ki a kéziratból, és végtelen türelemmel viselte az utolsó heteket is.

Irodalomjegyzék

- [1] W.P. Rowe, R.J. Huebner, L.K. Gilmore, R.H. Parrott, and T.G. Ward. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.*, pages 570–573, 1953.
- [2] F.M. Bennett, B.B. Law, W. Hamilton, and A. MacDonald. Adenovirus eye infections in Aberdeen. *Lancet*, 2:670–673, 1957.
- [3] T. Shinozaki, K. Araki, H. Ushijima, and R. Fujii. Antibody response to enteric adenovirus types 40 and 41 in sera from people in various age groups. *J. Clin. Microbiol.*, 25, 1987.
- [4] J.J. Trentin, Y. Yabe, and G. Taylor. The quest for human cancer viruses. *Science*, 137:835–849, 1962.
- [5] L.T. Chow, R.E. Gelinas, T.R. Broker, and R.J. Roberts. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12:1–8, 1977.
- [6] A.E. Smith. Viral vectors in gene therapy. *Annu. Rev. Microbiol.*, 49:807–838, 1995.
- [7] E. Vigne, J.F. Dedieu, C. Orsini, M. Latta, B. Klonjowski, E. Prost, M.M. Lakich, E.J. Kremer, P. Deneffe, M. Perricaudet, and P. Yeh. Improvement of adenoviral vectors for human gene therapy. In K. Berg, V. Boulyjenkov, and Y. Christen, editors, *Genetic Approaches to Non-communicable Diseases*, pages 113–131. Springer Verlag, 1996.
- [8] W.C. Russell. Update on adenovirus and its vectors. *J. Gen. Virol.*, 81:2573–2604, 2000.
- [9] M.P. Van den Ende, D. Kipps, and A. Kipps. The isolation in an eggs of a new filtrable agent which may be the cause of bovine lumpy skin disease. *J. Gen. Microbiol.*, 3:174–183, 1949.

- [10] V.J. Yates and D.E. Fry. Observation on a chicken embryo lethal orphan (CELO) virus. *Am. J. Vet. Res.*, 18:657–660, 1957.
- [11] M. Klein, E. Early, and J. Zellat. Isolation from cattle of a virus related to human adenoviruses. *Proc. Soc. Exp. Biol. Med.*, 102:1–4, 1959.
- [12] A Bartha. Proposal for subgrouping of bovine adenoviruses. *Acta Vet. Hung.*, 19:319–321, 1969.
- [13] E. Norrby, A. Bartha, P. Boulanger, R. S. Dreizin, H. S. Ginsberg, S. S. Kalter, H. Kawamura, W. P. Rowe, W. C. Russell, W. Schlesinger, and R. Wigand. Adenoviridae. *Interviol.*, 7:117–125, 1976.
- [14] F. Bürki, B. Hinaidy, and J. Böckman. Bovine adenoviruses III. sub-classification into two subgroups by means of immunodiffusion test. *Microbiologica*, 2:65–74, 1979.
- [15] R.P. Hendrick, J. Spears, M.L. Kent, and T. McDowell. Adenolike virus associated with a disease of cultured white sturgeon *Acipenser transmontanus*. *Can. J. Fisheries Aquatic Sci.*, 42:1231–1235, 1985.
- [16] E.R. Jacobson, C.H. Gardiner, and C.M. Foggin. Adenovirus like infection in two Nile crocodiles. *J. Am. Vet. Med. Assoc.*, 185:1421–1, 1985.
- [17] R.J. Roberts, G. Akusjärvi, P. Aleström, R.E. Gelinas, T.R. Gingeras, D. Sciaky, and U. Pettersson. A consensus sequence for the adenovirus genome. In W Doerfler, editor, *Adenovirus DNA*, pages 1–51. Boston: Martinus Nijhoff, 1986.
- [18] S. Vрати, D.E. Brookes, P. Strike, A. Khatri, D. B. Boyle, and G.W. Both. Unique genome arrangement of an ovine adenovirus: Identification of new proteins and proteinase cleavage sites. *Virology*, 220:186–199, 1996.
- [19] S. Chiocca, R. Kurzbauer, G. Schaffner, A. Baker, V. Mautner, and M. Cotten. The complete DNA sequence and genomic organization of the avian adenovirus CELO. *J. Virol.*, 70:2939–2949, 1996.
- [20] J. Pitcovski, M. Mualem, Z. Rei-Koren, S. Krispel, E. Shmueli, Y. Peretz, B. Gutter, G. E. Gallili, A. Michael, and D. Goldberg. The complete DNA sequence and genome organization of the avian adenovirus, hemorrhagic enteritis virus. *Virology*, 249:307–315, 1998.

- [21] M.D. Morrison, D.E. Onions, and L. Nicolson. Complete DNA sequence of canine adenovirus type 1. *J. Gen. Virol.*, 78:873–878, 1997.
- [22] M.J. Raaij, A. Mitraki, G. Lavigne, and S. Cusac. A triple β -spiral in the adenovirus fibre reveals a new structural motif for fibrous protein. *Nature*, 401:935–938, 1999.
- [23] U. Pettersson and S. Hoglund. Structural proteins of adenoviruses. *Virology*, 39:90–106, 1969.
- [24] S. Belák, Gy. Berencsi, M. Rusvai, K. Lukács, and I. Nász. DNA structure, and hemagglutination properties of bovine adenovirus type 2 strains which bypass species specificity. *Arch. Virol.*, 77:181–194, 1983.
- [25] S. Belák, A. Virtanen, J. Zabielski, M. Rusvai, Gy. Berencsi, and U. Pettersson. Subtypes of bovine adenovirus type 2 exhibit major differences in region E3. *Virology*, 153:262–271, 1986.
- [26] U.F. Greber, P. Webster, J. Weber, and A. Helenius. The role of the adenovirus protease in virus entry into cells. *EMBO J.*, 15:1766–1777, 1996.
- [27] U.F. Greber and A. Helenius. Mechanisms of virus uncoating. *Trends Microbiol.*, 2:52–56, 1994.
- [28] U.F. Greber, M. Willetts, P. Webster, and A. Helenius. Stepwise dismantling of adenovirus 2 during entry into cells. *Cell*, 75:447–486, 1993.
- [29] D.A. Matthews and W.C. Russell. Adenovirus core protein V interacts with p32 – a protein which is associated with both the mitochondria and the nucleus. *J. Gen. Virol.*, 79:1677–1685, 1998.
- [30] K. Hayashi and W.C. Russell. A study of the development of adenovirus antigens by the immunofluorescent technique. *Virology*, 34:470–480, 1968.
- [31] P. Lutz, F. Puvion-Dutilleul, Y. Lutz, and C. Kedinger. Nucleoplasmic and nucleolar distribution of the adenovirus iv_{a2} gene product. *J. Virol.*, 70:3449–3460, 1996.
- [32] D.A. Matthews and W.C. Russell. Adenovirus core protein V is delivered by the invading virus to the nucleus of the infected cell and later in infection is associated with nucleoli. *J. Gen. Virol.*, 1998:1671–1675, 1998.

- [33] L.R. Gooding and W.S.M. Wold. Molecular mechanisms by which adenoviruses counteract antiviral immune defense. *Immunology*, 10:53–71, 1990.
- [34] W.S.M. Wold, C. Cladaras, S.L. Deutscher, and Q.S. Kapoor. The 19-kDa glycoprotein coded by region E3 of adenovirus: purification, characterization, and structural analysis. *J. Biol. Chem.*, 260:2424, 1985.
- [35] S. Kvist, L. Östberg, H. Persson, L. Philipson, and P.A. Peterson. Molecular association between transplantation antigens and cell surface antigen in adenovirus-transformed cell line. *Proc. Natl. Acad. Sci. USA*, 75:5674, 1978.
- [36] S. Paabo, F. Weber, T. Nilsson, W. Schaffner, and P.A. Peterson. Structural and functional dissection of an MHC class I antigen-binding adenovirus glycoprotein. *EMBO J.*, 5:1921, 1987.
- [37] L.R. Gooding, L.W. Elmore, A.E. Tollefson, H.A. Brady, and W.S.M. Wold. A 14,700MW protein from E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell*, 53:341, 1988.
- [38] A.E. Tollefson, S.J. Ryerse, A. Scaria, T.W. Hermiston, and S.M. Wold. The E3–11.6-kDa adenovirus death protein (ADP) is required for efficient cell death: characterization of cells infected with *adp* mutants. *Virology*, 220:152–162, 1996.
- [39] D.N. Halbert, J.R. Cutt, and T. Shenk. Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *J. Virol.*, 56:250–257, 1985.
- [40] E. Bridge and G. Ketner. Redundant control of adenovirus late gene expression. *J. Virol.*, 63:631–638, 1989.
- [41] M. Matthews and T. Shenk. Adenovirus virus-associated RNA and translation control. *J. Virol.*, 65:5657–5662, 1991.
- [42] R. O’Malley, R. Duncan, J. Hershey, and M. Matthews. Modification of protein synthesis initiation factors and the shut-off of host protein synthesis in adenovirus-infected cells. *Virology*, 168:112–118, 1989.
- [43] K.N. Leppard. Regulated RNA processing and RNA transport during adenovirus infection. *Sem. Virol.*, 8:301–307, 1998.

- [44] R.A. Guggenheimer, K. Nagata, and J. Field. *In vitro* synthesis of full length adenoviral DNA. In N. Cozzarelli, editor, *UCLA Symposia on Mechanisms of DNA Replication and Recombination*, volume 10, pages 395–421. New York: Alan R. Liss, 1983.
- [45] J.F. Diffley and B. Stillman. Purification of a cellular double stranded DNA-binding protein required for initiation of adenovirus DNA replication by using a rapid filter binding assay. *Mol. Cell. Biol.*, 6:1363–1373, 1986.
- [46] B.R. Friefeld, J. Lichy, and J. Field. The *in vitro* replication of adenovirus DNA. *Curr. Top. Microbiol. Immunol.*, 110:221–255, 1984.
- [47] P. Wildy. Adenoviruses. In *Monogr. Virol.*, volume 5. Karger, Basel, 1971.
- [48] R. Wigand, A. Bartha, R.S. Dreizin, H. Esche, H.S. Ginsberg, M. Green, J.C. Hierholzer, S.S. Kalter, J.B. McFerran, U. Pettersson, W.C. Russell, and G. Wadell. *Adenoviridae*: Second Report. *Intervirology*, 18:169–176, 1982.
- [49] J. Hierholzer, A.C. Torrence, and P.R. Wright. Generalized viral illness caused by an intermediate strain of adenovirus 21/H21+31. *J. Infect. Dis.*, 14:281–288, 1980.
- [50] M. Nagy, É. Nagy, and T. Tuboly. Sequence analysis of porcine adenovirus serotype 5 fiber gene: Evidence for recombination. *Virus Genes*, 24:177–181, 2002.
- [51] M. Benkő, B. Harrach, and W.C. Russell. Family *Adenoviridae*. In M.H.V. van Regenmortel, C.M. Fauquet, D.H.L. Bishop, E.B. Carstens, M.K. Estes, S.M. Lemon, J. Maniloff, M.A. Mayo, D.J. McGeoch, C.R. Pringle, and R.B. Wickner, editors, *Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses*, pages 227–238. Academic Press, San Diego, 2000.
- [52] A. J. Davison and B. Harrach. Genus *Siadenovirus*. In C.A. Tidona and G. Darai, editors, *The Springer Index of Viruses*, pages 29–33. Springer-Verlag, Heidelberg, 2002.
- [53] H.D. Lehmkuhl and R.C. Cutlip. A new goat adenovirus isolate proposed as the prototype strain for goat adenovirus serotype 1. *Arch. Virol.*, 1611–1618:144, 1999.

- [54] A.J. Davison, K.M. Wright, and B. Harrach. DNA sequence of frog adenovirus. *J. Gen. Virol.*, 81:2431–2439, 2000.
- [55] A. N. Zakharchuk, V. A. Kruglyak, T. A. Akopian, and T. I. Tikhonenko. Physical mapping and homology studies of the egg drop syndrome (EDS-76) adenovirus DNA. *Arch. Virol.*, 128:171–176, 1993.
- [56] R.M. Adair, E.R. McKillop, and B.H. Coakley. Serological identification of an australian adenovirus isolate from sheep. *Aust. Vet. J.*, 63:162, 1986.
- [57] M. Hess, H. Blöcker, and P. Brandt. The complete nucleotide sequence of the egg drop syndrome virus: an intermediate between mastadenoviruses and aviadenoviruses. *Virology*, 238:145–156, 1997.
- [58] Harrach. B. Az adenovírusok filogenetikája, MTA doktori tézisek, Budapest, 2001.
- [59] M. Benkő, A. Bartha, and G. Wadell. DNA restriction enzyme analysis of bovine adenoviruses. *Intervirol.*, 29:346–350, 1988.
- [60] M.J. Short, J.M. Fernandez, J.A. Sorge, and W.D. Huse. ZAP: A λ bacteriophage expression vector with *in vivo* excision properties. *Nucleic Acids Res.*, 16:7583–7600, 1988.
- [61] J. Sambrook, E.F. Fritsch, and T. Maniatis. *Molecular Cloning*. Cold Spring Harbor Laboratory Press, 1989.
- [62] A. Nishimura, M. Morita, Y. Nishimura, and Y. Sugino. A rapid and highly efficient method for preparation of competent *Escherichia coli* cells. *Nucl. Acids Res.*, 18:6169, 1990.
- [63] R. Staden. The Staden sequence analysis package. *Mol. Biotech.*, 5:233–241, 1996.
- [64] Stephen F. Altschul, Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.*, 25:3389–3402, 1997.
- [65] Timothy L. Bailey and C. Elkan. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. In *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, pages 28–36. AAAI Press, Menlo Park, California, 1994.

- [66] Timothy L. Bailey and M. Gribskov. Combining evidence using p-values: application to sequence homology searches. *Bioinformatics*, 14:48–54, 1998.
- [67] S. Karlin and S.F. Altschul. Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl. Acad. Sci. USA*, 90:5873–5877, 1993.
- [68] B. Rost. Predicting one-dimensional protein structure by profile based neural networks. *Methods Enzymol.*, 266:525–539, 1996.
- [69] J.D. Thompson, D.G. Higgins, and T.J. Gibson. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acid Res.*, 22:4673–4680, 1994.
- [70] G. J. Barton. ALSCRIPT A tool to format multiple sequence alignments. *Prot. Eng.*, 6:37–40, 1993.
- [71] J. Felsenstein. PHYLIP-Phylogeny inference package. *Cladistics*, 5:164–166, 1989.
- [72] R.D.M. Page. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comp. Appl. in Biosci.*, 12:357–358, 1996.
- [73] Yang Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS*, 13:555–556, 1997.
- [74] Z. Yang. *Phylogenetic analysis by maximum likelihood (PAML), Version 3.0*. University College London, 2000.
- [75] M.O. Dayhoff, R.M. Schwartz, and B.C. Orcutt. A model of evolutionary change in proteins. In M.O. Dayhoff, editor, *Atlas of Protein sequence and Structure*, volume 5, pages 345–352. National Biochemical Research Foundation, Washington, D.C., 1978.
- [76] B. Harrach and M. Benkő. Phylogenetic analysis of adenovirus sequences: proof of the necessity of establishing a third genus in the adenoviridae family. In W.S.M. Wold and N.J. Totowa, editors, *Methods in Molecular Medicine*, volume 21, pages 309–339. Humana Press, 1998.
- [77] A. Khatri and G.W. Both. Identification of transcripts and promoter regions of ovine adenovirus OAV287. *Virology*, 245:128–141, 1998.

- [78] L. Fitzgerald, M. Ansari, and Y. Haj-Ahmad. Cloning and sequencing of the bovine adenovirus type 2 early region 4. *Gene*, 185:181–186, 1997.
- [79] M. Sheppard and K.M. Erny. Dna sequence analysis of the inverted terminal repeats of a non-oncogenic avian adenovirus. *Nucl. Acid Res.*, 17:3995, 1989.
- [80] R.J. Wides, M.D. Challberg, D.R. Rawlins, and T.J. Kelly. Adenovirus origin of replication:sequence requirements for replication *in vitro*. *Mol. Cell. Biol.*, 7:864–874, 1987.
- [81] M.D. Challberg and D.R. Rawlins. Template requirements for the initiation of adenovirus DNA replication. *Proc. Natl. Acad. Sci. USA*, 81:100–104, 1984.
- [82] M.P. Harris and R.T. Hay. DNA sequences required for the initiation of adenovirus type 4 DNA replication *in vitro*. *J. Mol. Biol.*, 201:57–67, 1988.
- [83] E. de Vries, W. van Driel, M. Tromp, J. van Boom, and P.C. van der Vliet. Adenovirus DNA replication *in vitro*: site-directed mutagenesis of the nuclear factor I binding site of the Ad2 origin. *Nucl. Acid Res.*, 13:4935–4952, 1985.
- [84] P.A. Leegwater, W. van Driel, and P.C. van der Vliet. Recognition site of nuclear factor I, a sequence-specific DNA-binding protein from HeLa cells that stimulates adenovirus DNA replication. *EMBO J.*, 4:1515–1521, 1985.
- [85] G. J. Pruijn, R.T. van Miltenburg, J. A. Claessens, and P.C. van der Vliet. Interaction between the octamer-binding protein nuclear factor III and the adenovirus origin of DNA replication. *J. Virol.*, 62:3092–3102, 1988.
- [86] L. Hatfield and P. Hearing. Redundant elements in the adenovirus type 5 inverted terminal repeat promote bidirectional transcription *in vitro* and are important for virus growth *in vivo*. *Virology*, 184:265–276, 1991.
- [87] Y.M. Mul, C.P. Verrijzer, and P.C. van der Vliet. Transcription factors NFI and NFIII/OCT-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. *J. Virol.*, 64:5510–5518, 1990.

- [88] L. Hatfield and P. Hearing. The NFIII/OCT-1 binding site stimulates adenovirus DNA replication *in vivo* and is functionally redundant with adjacent sequences. *J. Virol.*, 67:3931–3939, 1993.
- [89] M. Shinagawa, T. Ishiyama, R. Padmanabhan, K. Fujinaga, M. Kamada, and G. Sato. Comparative sequence analysis of the inverted terminal repetition in the genomes of animal and avian adenoviruses. *Virology*, 125:491–495, 1983.
- [90] M. Shinagawa, Y. Iida, A. Matsuda, T. Tsukiyama, and G. Sato. Phylogenetic relationships between adenoviruses as inferred from nucleotide sequences of inverted terminal repeats. *Gene*, 1987.
- [91] S.M. Temperley, C.R. Burrow, T.J. Kelly, and R.T. Hay. Identification of two distinct regions within the adenovirus minimal origin of replication that are required for adenovirus type 4 DNA replication *in vitro*. *J. Virol.*, 65:5037–5044, 1991.
- [92] Y. Shen, G. Kitzes, J.A. Nye, A. Fattaey, and T. Hermiston. Analysis of single amino acid substitution mutants of adenovirus type 5 E1B 55K protein. *J. Virol.*, 75:4297–4307, 2001.
- [93] J.A. Engler, M.S. Hoppe, and M.P. van Bree. The nucleotide sequence of the genes encoded in early region 2b of human adenovirus type 7. *Gene*, 21:145–159, 1983.
- [94] T.R. Gingeras, D. Sciaky, R.E. Gelinis, J. Bing-Dong, C.E. Yen, M.M. Kelly, P.A. Bullock, B.L. Parsons, K.E. O’Neill, and R.J. Roberts. Nucleotide sequences from the adenovirus-2 genome. *J. Biol. Chem.*, 257:3475–3491, 1982.
- [95] A. Khatri and G.W. Both. Identification of transcripts and promoter regions of ovine adenovirus OAV287. *Virology*, 245:128–141, 1998.
- [96] J.L. Boyer and G. Ketner. Genetic analysis of a potential zinc-binding domain of the adenovirus E4 34k protein. *J. Biol. Chem*, 275:14969–14978, 2000.
- [97] J.S. Orlando and D.A. Ornelles. E4orf6 variants with separate abilities to augment adenovirus replication and direct nuclear localization of the e1b 55-kilodalton protein. *J. Virol.*, 76:1475–87, 2002.
- [98] E. Querido, P. Blanchette, T. Yan, Q. Kamura, D. Morrison, M. Boivin, W.G. Kaelin, R.C. Conaway, J.W. Conaway, and E. Branton, P.

- Degradation of p53 by adenovirus E4orf6 and E1B55K proteins occurs via a novel mechanism involving a Cullin-containing complex. *Genes Dev.*, 15:3104–17, 2001.
- [99] T. Dosch, F. Horn, G. Schneider, F. Kratzer, T. Dobner, J. Hauber, and R.H. Stauber. The adenovirus type 5 E1B-55K oncoprotein actively shuttles in virus-infected cells, whereas transport of E4orf6 is mediated by a CRM1-independent mechanism. *J. Virol.*, 75:5677–83, 2001.
 - [100] P.A. Eagle and D.F. Klessig. A zinc-binding motif located between amino acids 273 and 286 in the adenovirus DNA-binding protein is necessary for ssDNA binding. *Virology*, 187:777–787, 1992.
 - [101] S. Vрати, D.E. Brookes, D.B. Boyle, and G.W. Both. Nucleotide sequence of ovine adenovirus tripartite leader sequence and homologues of the IVa2, DNA polymerase and terminal proteins. *Gene*, 177:35–41, 1996.
 - [102] T. B. Hasson, D.A. Ornelles, and T. Shenk. Adenovirus 11 52- and 55-kilodalton proteins are present within assembling virions and colocalize with nuclear structures distinct from replication centers. *J. Virol.*, 66:6133–6142, 1992.
 - [103] J.M. Weber and K. Tihanyi. Adenovirus endopeptidases. *Methods Enzimol.*, 244:595–604, 1994.
 - [104] B. Harrach, B.M. Meehan, M. Benkő, B.M. Adair, and D. Todd. Close phylogenetic relationship between egg drop syndrome virus, bovine adenovirus serotype 7, and ovine adenovirus strain 287. *Virology*, 229:302–308, 1997.
 - [105] R. Szathmáry, C. Barbezange, Á. Dán, and B. Harrach. Simultaneous physical and genetic mapping of viruses: cloning, sequencing, and mapping of the genome of bovine adenovirus type 6. In *NADC First International Virtual Conference on the Infectious Disease of Animals*, number B00011 in 0. NADC, Ames, Iowa, 1997.
 - [106] C. Sander; R. Schneider. Database of homology-derived structures and the structural meaning of sequence alignment. *Proteins*, 9:56–68, 1991.
 - [107] I. Bagyan, B. Setlow, and P. Setlow. New small, acid-soluble proteins unique to spores of *Bacillus subtilis*: identification of the coding genes and regulation and function of two of these genes. *J. Bacteriol.*, 180:6704–6712, 1998.

- [108] E.R. Fliss and P. Setlow. Genes for *Bacillus megaterium* small, acid-soluble spore proteins: nucleotide sequence of two genes and their expression during sporulation. *Gene*, 35:151–157, 1985.
- [109] N.G. Magill, C.A. Loshon, and P. Setlow. Small, acid-soluble, spore proteins and their genes from two species of *Sporosarcina*. *FEMS Microbiol. Lett.*, 60:293–297, 1990.
- [110] F. Kunst, N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S.C. Brignell, S. Bron, S. Brouillet, C.V. Bruschi, B. Caldwell, V. Capuano, N.M. Carter, S.K. Choi, J.J. Codani, I.F. Connerton, and A. Danchin. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature*, 390:249–256, 1997.
- [111] M.J. Connors and P. Setlow. Cloning of a small, acid-soluble spore protein gene from *Bacillus subtilis* and determination of its complete nucleotide sequence. *J. Bacteriol.*, 161:333–339, 1985.
- [112] S. Henikoff, J.G. Henikoff, W.J. Alford, and S. Pietrokovski. Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene*, 163:17–26, 1995.
- [113] S. D. Benson, J.K. Bamford, D.H. Bamford, and R.M. Burnett. Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell*, 98:825–833, 1999.
- [114] I. Kiss, K. Matiz, A. Allard, G. Wadell, and M Benkő. Detection of homologous DNA sequences in animal adenoviruses by polymerase chain reaction. *Acta Vet. Hung.*, 44:243–251, 1996.
- [115] A. Allard, B. Albinsson, and G. Wadell. Rapid typing of human adenoviruses by a general pcr combined with restriction endonuclease analysis. *J. Clin. Microbiol.*, 39:498–50, 2001.
- [116] K. Ganesh, V.V. Suryanarayana, and R. Raghavan. Detection of fowl adenovirus associated with hydropericardium hepatitis syndrome by a polymerase chain reaction. *Vet. Res. Commun.*, 26:73–80, 2002.
- [117] A. Abraham, D. Rattner, and A. Zissman. Isolation of an adenovirus from the goatex – the wild ibex – domestic gat hybrid. *Isr. J. Vet. Med.*, 45:29–31, 1989.

- [118] M.M. Roberts, J.L. White, M.G. Grutter, and R.M. Burnett. Three-dimensional structure of the adenovirus major coat protein hexon. *Science*, 232:1148–1151, 1986.
- [119] C.I. Toogood and R.T. Hay. DNA sequence of the adenovirus type 41 hexon gene and predicted structure of the protein. *J. Gen. Virol.*, 69:2291–2301, 1988.
- [120] L. Crawford-Miksza and D.P. Schnurr. Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. *J. Virol.*, 70:1836–1844, 1996.
- [121] Á. Dán, Zs. Ruzsics, W.C. Russell, M. Benkő, and B. Harrach. Analysis of the hexon gene sequence of bovine adenovirus type 4 provides further support for a new adenovirus genus (*Atadenovirus*). *J. Gen. Virol.*, 79:1453–1460, 1998.
- [122] M.L. Boudin, J.C. D’Halluin, C. Cousin, and P. Boulanger. Human adenovirus type 2 protein IIIa. *Virology*, 101:144–156, 1980.
- [123] P. Mathias, T. Wickham, M. Moore, and G. Nemerow. Multiple adenovirus serotypes use alpha v integrins for infection. *J. Virol.*, 68:6811–6814, 1994.
- [124] A.J. Davison, E.A. Telford, M.S. Watson, K. McBride, and V. Mautner. The DNA sequence of adenovirus type 40. *J. Mol. Biol.*, 234:1308–1316, 1993.
- [125] J.D. Meissner, G.N. Hirsch, E.A. LaRue, R.A. Fulcher, and K.R. Spindler. Completion of the DNA sequence of mouse adenovirus type 1: sequence of E2B, L1, and L2. *Virus Res.*, 51:53–64, 1997.
- [126] P.S. Reddy, N. Idamakanti, J.Y. Song, J.B. Lee, B.H. Hyun, J. H. Park, S. H. Cha, Y.T. Bae, S. K. Tikoo, and L.A. Babiuk. Nucleotide sequence and transcription map of porcine adenovirus type 3. *Virology*, 251:414–426, 1998.
- [127] P.S. Reddy, N. Idamakanti, A.N. Zakhartchouk, M.K. Baxi, J.B. Lee, C. Pyne, L.A. Babiuk, and S.K. Tikoo. Nucleotide sequence, genome organization, and transcription map of bovine adenovirus type 3. *J. Virol.*, 72:1394–1402, 1998b.

- [128] A. Webster, R.T. Hay, and G. Kemp. The adenovirus protease is activated by a virus-encoded disulphide-linked peptide. *Cell*, 72:97–104, 1993.
- [129] M. Rusvai, B. Harrach, A. Bánrévi, P. S. Evans, and M. Benkő. Identification and sequence analysis of the core protein genes of bovine adenovirus 2. *Virus Res.*, 70:25–30, 2000.
- [130] M. Benkő and B. Harrach. A proposal for a new (third) genus within the *Adenoviridae* family. *Arch. Virol.*, 143:829–837, 1998.

A. Függelék

Az illesztésekhez használt fehérjék génbanki elérhetősége

Vírus	Gén vagy Fehérje	Elérési szám
BAdV-1	ITR	M17112
BAdV-2	DBP	AAB16762
BAdV-2	proteáz	U44124
BAdV-3	hexon	K01264
BAdV-3	ITR	M17113
BAdV-3	E1B55K	Q64844
BAdV-3	proteáz	X53990
BAdV-4	DBP	Q997H7
BAdV-4	E4.2	Q997H0
BAdV-4	hexon	AF036092
BAdV-4	ITR	AF238883
BAdV-4	LH2	Q997I7
BAdV-4	LH3	Q997I6
BAdV-4	fiber	Q997H2
BAdV-4	p32K	NP_077384
BAdV-4	RH2	Q997G4
BAdV-4	RH6	Q997G8
BAdV-4	RH1	Q997G3
BAdV-4	RH4	Q997G6
BAdV-5	hexon	AF207658
BAdV-5	ITR	AF238881
BAdV-6	hexon	AF2076559
BAdV-7	hexon	AF238232
BAdV-7	proteáz	X53989

Vírus	Gén vagy Fehérje	Elérési szám
BAdV-8	hexon	AF238233
BAdV-10	DBP	AAC35886
BAdV-10	ITR	AF238882
BAdV-10	proteáz	AF027599
CAdV-1	DBP	AAB05445
CAdV-1	E4 34K	Q96690
CAdV-1	hexon	U55001
CAdV-1	ITR	Y07760
CAdV-2	E4 34K	P87568
CAdV-2	hexon	U77082
CAdV-2	ITR	U77082
DAdV-1	19.4k	O11427
DAdV-1	19.9k	O11427
DAdV-1	DBP	O11422
DAdV-1	E1BS	O11412
DAdV-1	E4	O11425
DAdV-1	fiber	O11424
DAdV-1	hexon	Y09598
DAdV-1	ITR	Z86066
DAdV-1	LH3	O11412
DAdV-1	p24K	CAA70797
EAdV-1	hexon	L79955
EAdV-1	ITR	K009377
EAdV-2	hexon	L80007
FAdV-1	DBP	AAC54914
FAdV-1	hexon	U46993
FAdV-1	ITR	K00940
FAdV-8	DBP	Q9QM71
FAdV-9	hexon	AF083975
FAdV-9	ITR	AF02125310
FAdV-10	hexon	U26221
FrAdV-1	DBP	NP_062445
FrAdV-1	hexon	AF224336
FrAdV-1	ITR	AF224336
GAdV-1	hexon	AF207660
HAdV-1	hexon	AF161519
HAdV-2	DBP	NP_040527
HAdV-2	hexon	J01917
HAdV-2	E4 34K	P03239

Vírus	Gén vagy Fehérje	Elérési szám
HAdV-2	LH2	P87658
HAdV-2	ITR	J01917
HAdV-3	hexon	X76549
HAdV-3	ITR	J01960
HAdV-3	proteáz	X13271
HAdV-4	hexon	X84646
HAdV-4	ITR	J01964
HAdV-4	DBP	P06500
HAdV-4	E1B 55K	P03243
HAdV-5	DBP	P03265
HAdV-5	fiber	P11818
HAdV-5	hexon	M73260
HAdV-6	hexon	X67710
HAdV-7	DBP	P04497
HAdV-7	hexon	X76551
HAdV-7	E1B 55K	P03245
HAdV-8	hexon	AF161561
HAdV-9	E4 34K	P89083
HAdV-9	hexon	AF161562
HAdV-9	ITR	J01994
HAdV-12	hexon	X73487
HAdV-12	ITR	X73487
HAdV-12	DBP	P04498
HAdV-16	hexon	X74662
HAdV-17	hexon	AF108105
HAdV-19	hexon	AF161565
HAdV-21	hexon	AJ012091
HAdV-37	hexon	AB023555
HAdV-40	hexon	X51782
HAdV-41	DBP	AAA52196
HAdV-41	hexon	X51783
HAdV-41	ITR	M20300
HAdV-48	hexon	U20821
MAdV-1	hexon	U57336
MAdV-1	ITR	M37187
MAdV-1	proteáz	M19540
OAdV-3	proteáz	AF153447
OAV287	DBP	Q83907
OAV287	E4.3	P88027

Vírus	Gén vagy Fehérje	Elérési szám
OAV287	E4.2	P88027
OAV287	hexon	U40837
OAV287	ITR	U40839
OAV287	LH2	Q83923
OAV287	LH3	Q83924
OAV287	fiber	Q83893
OAV287	p32K	AAB19237
OAV287	RH1	Q83921
OAV287	RH2	Q83920
OAV287	RH4	Q83918
OAV287	RH6	Q83916
OdAdV-1	E1B T	Q9QCU5
OdAdV-1	hexon	AF198354
PAdV-1	ITR	L43075
PAdV-2	ITR	L43076
PAdV-3	DBP	CAB41032
PAdV-3	hexon	U34592
PAdV-3	ITR	L43077
PAdV-4	ITR	L43078
PAdV-5	DBP	Q994D7
PAdV-5	E4 ORF6	Q98VKA
PAdV-5	ITR	L43079
PoAdV-1	hexon	AF338822
PoAdV-1	proteáz	AF338823
Rus	ITR	AF238880
SAdV-8	ITR	M28838
SAdV-21	hexon	AR101858
SAdV-25	DBP	Q8UY78
SAdV-25	hexon	AR101859
TAdV-3	DBP	NC_001958
TAdV-3	hexon	AF074946
TAdV-3	ITR	AF074946
TsAdV-1	ITR	V01485
<i>Bacillus megaterium</i>	SAS2_BACME	P10571
<i>Bacillus subtilis</i>	SAS_BACSU	NP_390835
<i>Bacillus subtilis</i>	SAS3_BACSU	P02958
<i>Sporosarcina halophila</i>	SAS1_SPOHA	Q00213