

**DOKTORI ÉRTEKEZÉS**

*Fátyol Károly*

**EGY ÚJ KROMOSZÓMA SPECIFIKUS CENTROMER SZEKVENCIA  
AZONOSÍTÁSA ÉS MOLEKULÁRIS JELLEMZÉSE**

*Témavezető: Dr Hadlaczky Gyula*

*Szegedi Biológiai Központ*

1997

## TARTALOM

1. Irodalmi áttekintés.....	1
1.1. Az eukarióta genom reasszociációs kinetikája.....	2
1.2. Repetitív DNS szekvenciák.....	3
1.2.1. Szétszórt repetitív elemek.....	3
1.2.1.1. Retropozon eredetű repetitív elemek.....	4
1.2.1.1.1. Rövid szétszórt elemek.....	4
1.2.1.1.2. Hosszú szétszórt elemek.....	5
1.2.1.1.3. Retrovírus-szerű elemek.....	6
1.2.1.2. DNS transzpozon eredetű repetitív elemek.....	6
1.2.2. Tandem repetitív elemek.....	7
1.2.2.1. Tandem repetitív gén-családok.....	9
1.2.2.2. Szatellitek.....	10
1.2.2.3. Miniszatellitek ("egyszerű" szekvenciák).....	12
1.3. Repetitív szekvenciák és a kromoszóma szerkezet kapcsolata.....	13
1.3.1. DNS replikációs origó.....	14
1.3.2. Telomer.....	15
1.3.3. Centromer.....	17
1.3.3.1. A centromer-kinetokor komplex funkciói.....	18
1.3.3.2. A centromer-kinetokor komplex molekuláris felépítése.....	20
2. Célkitűzések.....	24
3. Anyagok és módszerek.....	25
3.1. Emlős sejtek tenyésztése.....	25

3.2. Kínai hörcsög kromoszómák <i>in vivo</i> fragmentációja, centromerikus kromoszóma darabok immunoprecipitációja.....	25
3.3. <i>In situ</i> hibridizáció.....	26
3.4. DNS technikák.....	28
3.5. Szekvencia analízis.....	28
3.6. <i>Pulsed field</i> gélelektroforézis.....	28
3.7. PCR, RT-PCR.....	29
4. Eredmények.....	30
4.1. Kromoszóma specifikus centromer szekvencia izolálása kínai hörcsög sejtéből.....	30
4.2. Az azonosított centromerikus DNS fragment egy új, faj es kromoszóma specifikus repetitív DNS elem része.....	33
4.3. Az "a" es "b" típusú ismétlődő egységek szekvencia analízise.....	36
4.3.1. Az ismétlődő egységek különleges DNS elemeket tartalmaznak.....	37
4.3.2. A "b" típusú repetitív elem valószínűleg gén konverzió eredménye.....	39
4.3.3. Az "a" es "b" elemek határán egy hosszú, kettős homeobox szakaszt tartalmazó nyílt leolvasási keret található.....	39
4.4. Az 1.2 kb méretű BamHI-SacI fragment egy deléciós repetitív egységet tartalmaz.....	42
4.5. Az új repetitív elem a 2. kínai hörcsög kromoszóma pericentrikus régiójának az egyik fő szatellit-szerű alkotó eleme.....	42
4.6. A HC2 szatellit nagy méretű homogén doméneket alkot.....	45
4.7. A HC2 szatellittel homológ DNS elemek fordított ismétlődéseket tartalmaznak.....	47
5. Összefoglalás.....	50
6. Summary.....	57
7. Irodalom.....	61

8. Mellékletek.....	66
1. Melléklet: Az "a" típusú ismétlődő egység nukleotid szekvenciája.....	66
2. Melléklet: A "b" típusú ismétlődő egység nukleotid szekvenciája.....	68
3. Melléklet: Az 1.2 kb méretű deléciós ismétlődő egység nukleotid szekvenciája.....	70
Köszönet nyilvánítás.....	71



## 1. IRODALMI ÁTTEKINTÉS

A haploid genomot alkotó teljes DNS mennyiség a fajra jellemző meghatározott érték (C-érték). A C-érték rendkívüli mértékű változatosságot mutat a törzsfajlás során. A skála a mycoplasmák kevesebb mint  $10^6$  bp méretű genomjától a kétéltűek és bizonyos növények  $10^{11}$  bp méretű genomjáig terjed. Általános tendenciaként megfigyelhető, hogy a testszerveződés komplexitásának növekedésével együtt jár a genom méretének növekedése. Az eukarióta sejtszerveződés kialakulása, a többsejtű életforma megjelenése elkerülhetetlenül új funkciók kialakulását eredményezte. Az ezen funkciókat végrehajtó fehérje és RNS molekulák kódolása szükségszerűen a genetikai információs rendszer bővülését tette szükségessé. A testszerveződés komplexitása és a genom mérete közötti kapcsolat azonban sok esetben nem egyértelmű. A kétéltűek haploid genomjának mérete  $\sim 8 \times 10^8$  bp-tól  $\sim 10^{11}$  bp-ig terjed. Az emlősök haploid genomja ugyanakkor csak  $3-4 \times 10^9$  bp. Szervezettenileg és törzsfajlásban az emlősök testszerveződése egyértelműen magasabb szintűnek tekinthető mint a kétéltűeké. A számos kétéltű fajban ennek ellenére megfigyelhető, az emlősökét felülmúló genom méret nem magyarázható új funkciók megjelenésével. Ez és számos hasonló példa képezi alapját az úgynevezett C-érték paradoxonnak. A C-érték paradoxon jelenlegi tudásunk azon hiányosságát tükrözi miszerint képtelenek vagyunk az eukarióta sejtmag DNS tartalmának többségéhez funkciót rendelni. Mivel a magasabb rendű eukarióták DNS-ének csak töredéke kódol fehérje vagy RNS molekulákat a kérdés megoldásához a genom nagyobb részét kitevő, nem kódoló DNS szakaszok vizsgálata vihet közelebb bennünket. Ennek a DNS frakciónak jelentős részét több kópiában előforduló ún. repetitív elemek alkotják.

Ebben a szakaszban először röviden ismertetem a magasabbrendű eukariótákból, elsősorban emlősből ismert repetitív DNS elemek szerveződését, majd ezen DNS elemek és a kromoszóma szerkezet és funkció kapcsolatát fogom áttekinteni.

### ***1.1. Az eukarióta genom reaszociációs-kinetikája:***

Az eukarióta genom DNS tartalmának összetételére a DNS reaszociációs-kinetikai vizsgálatai szolgáltatnak adatokat. Megfelelő fizikai, kémiai feltételek mellett a kettősszalú DNS molekulák komplementer szálai elválaszthatók egymástól (denaturáció). A fiziko-kémiai paraméterek megváltoztatásával a denaturációs folyamat megfordítható. A denaturáció fodítottjaként lejátszódó renaturációs folyamatok másodrendű kinetikát követnek. Adott paraméterek mellett a renaturáció sebessége arányos az egyszalú komplementer DNS molekulák koncentrációjának a négyzetével. A reaszociációt leíró sebességi egyenlet integrálásával majd megfelelő átrendezésével bevezethetjük a különböző DNS frakciók renaturációját jellemző úgynevezett  $Cot_{1/2}$  értéket, ami a komplementer egyszalú DNS molekulák koncentrációjának és az adott DNS frakció 50%-os renaturációjához szükséges időnek a szorzata. A fentebbi definícióból következik, hogy azonos DNS koncentrációk mellett nagyobb  $Cot_{1/2}$  érték lassabb, illetve kisebb  $Cot_{1/2}$  érték gyorsabb renaturációt jelent.

A reaszociálódott DNS frakció százalékát ábrázolva a  $Cot_{1/2}$  érték logaritmusának függvényében az úgynevezett Cot görbéhez jutunk. A prokarióta genom renaturációjának Cot görbéje jól közelíti egy ismétlődő szakaszokat nem tartalmazó, elméleti genom renaturációját leíró szigmoid alakú görbét. Az eukarióta genom több platót tartalmazó összetett Cot görbéje három különböző sebességgel reaszociálódó DNS frakció létezésére utal:

1., Gyorsan reaszociálódó komponens: ezt a teljes genom ~25%-át kitevő DNS frakciót túlnyomórészt rövid, egyszerű szekvencia részletek tandem ismétlődéséből felépülő ún. szatellit DNS alkotja.

2., Közepes sebességgel reasszociálódó komponens: a genom ~30%-át reprezentáló főleg szétszórt, transzpozonszerű ismétlődő elemeket, valamint több kópiás gén családokat (riboszomális RNS gének, hiszton gének, aktin gének, globin gének stb.) magába foglaló DNS frakció. Az 1. és 2. frakciót alkotó DNS szakaszokat összefoglalóan ismétlődő vagy repetitív elemeknek nevezzük.

3., Az egy példányban jelenlevő DNS elemek alkotják az eukarióta genom lassan reasszociálódó komponensét. Ez az úgynevezett nem repetitív DNS frakció a prokarióta genom egyetlen komponense. Eukariótákban ez a frakció a teljes genom ~45%-át alkotja.

### **1.2. Repetitív DNS szekvenciák:**

A repetitív DNS szekvenciákat szerveződésük alapján két nagy csoportba sorolhatjuk:

- 1., Szétszórt repetitív elemek (*interspersed repeated elements*)
- 2., Tandem repetitív szekvenciák

#### **1.2.1. Szétszórt repetitív elemek:**

A szétszórt repetitív elemek családjába a genomban többé-kevésbé egyenletesen eloszló, transzpozon eredetű DNS szekvenciák tartoznak. Egy korábbi csoportosítás szerint a szétszórt elemek eloszlása a genomban alapvetően kétféle mintázatot követ (1):

- (i) a *Drosophila* mintázatra a néhány ezer bázispár hosszúságú repetitív elemek a jellemzőek amiket több tíz kilobázispárnyi nem ismétlődő DNS szakaszok választanak el,
- (ii) a *Xenopus* mintázatot a nagyobb sűrűségben előforduló néhány száz bázispár hosszúságú repetitív elemek dominanciája jellemzi. A szétszórt ismétlődő elemek szerveződése a vizsgált fajok többségében, beleértve az emlősöket is a *Xenopus* mintázatot mutatja.

A eukarióta genomban előforduló szétszórt DNS elemek közös sajátága a mobilis genetikai elemekre visszavezethető eredetük. A mobilis genetikai elemek terjedésük mechanizmusa alapján kétfélék lehetnek. A retropozonok RNS intermediéren keresztül



reverz transzkripcióval terjednek a genomon belül. A transzpozíció nélkülözhetetlen feltétele a transzpozon által kódolt RNS függő DNS polimeráz (reverz transzkriptáz). A DNS transzpozonok terjedése RNS intermediertől függetlenül történik. Az elem kivágódását és új helyre történő integrálódását az ún. transzpozáz katalizálja.

Mind a retropozonok, mind a DNS transzpozonok esetében előfordulhat, hogy mutációk inaktiválják az elem mobilizálásához szükséges enzimet. Ezek az elemek többé már nem képesek önállóan mozogni a genomban. Az ilyen mutáns elemek csak intakt ún. autonóm elemek jelenlétében képesek mozogni. Az autonóm elem által kódolt transzpozíciót katalizáló enzim nemcsak magát az intakt elemet, hanem az enzim cél szekvenciáját még tartalmazó mutáns kópiákat is mobilizálhatja. Az emlősök transzpozon eredetű repetitív elemei többségükben önálló mozgásra már képtelen mutáns DNS elemek.

#### **1.2.1.1. Retropozon eredetű repetitív elemek:**

A retropozon eredetű repetitív elemek közé három nagy szekvencia család tartozik: a rövid szétszórt elemek (SINE: *short interspersed nuclear element*), a hosszú szétszórt elemek (LINE: *long interspersed nuclear element*), és a retrovírus-szerű elemek (RLE: *retrovirus-like element*) (2).

##### **1.2.1.1.1. Rövid szétszórt elemek:**

A SINE elemek RNS polimeráz III által átírt RNS molekulákra vezethetők vissza. Közös jellemzőjük a rövid, néhány száz bázispárnyi méret, a belső RNS polimeráz III promóter, a 3' végi poli-A szakasz és az elemet határoló, az integráció cél szekvenciájának megfelelő rövid direkt ismétlődések. Legismertebb képviselőjük a humán genomban több mint egy millió példányban megtalálható ~300 bp méretű *Alu* elem. Az *Alu* elem szerkezetét tekintve egy 130 bp méretű DNS darab tandem dimerjének tekinthető, ahol a jobb oldali monomer egy 31 bp méretű integrálódott DNS szakaszt tartalmaz. A monomer 5' vég felőli 90 bázispárnyi szakasza nagyfokú szekvencia homológiát mutat a szignál peptidet felismerő

partikulum 7 SL RNS-ével. Az *Alu* elem konszenzus szekvenciája ugyancsak tartalmaz egy, az RNS polimeráz III promóterével homológ DNS szakaszt. Az egér és kínai hörcsög genomban megtalálható 130 bp méretű B1 elemek ekvivalensnek tekinthetők az *Alu* elemet alkotó monomerrel. A SINE elemek közé tartoznak még tRNS és snRNS eredetű szekvenciák. A tRNS eredetű elemek legelterjedtebb képviselői az emlősök többségében előforduló MIR elemek (*mammalian-wide interspersed repeat*) (3), míg például a selyemlepkében megtalálható Bm1 szekvenciák U1 snRNS eredetűek (4). A SINE elemek mobilitásának kérdése jelenleg sem teljesen tisztázott. Az RNS-ről DNS kópiát átíró reverz transzkriptáz enzim forrásául valószínűleg LINE elemek szolgálnak. LINE elemek 3' vége és a bizonyos SINE elemek között megfigyelt szekvencia homológiák bizonyítani látszanak ezt a feltételezést (5).

#### **1.2.1.1.2. Hosszú szétszórt elemek:**

A LINE és RLE elemek a szétszórt repetitív szekvenciák virális csoportját alkotják. Közös sajátosságuk, hogy ép kópiáik nyílt leolvasási kereteket tartalmaznak amelyek közül az egyik reverz transzkriptázt kódol. Az emlősök genomja  $2 \times 10^4$ - $1 \times 10^5$  LINE (L1) elemet tartalmaz. A LINE elemek konszenzus szekvenciája ~6500 bp hosszú, két nyílt leolvasási keretet tartalmaz és 3' vége poli-A szakaszban végződik. A két nyílt leolvasási keret közül az egyik RNS kötő fehérjét a másik pedig egy endonukleáz és reverz transzkriptáz aktivitással rendelkező enzimet kódol. Az utóbbi enzim kettős katalitikus aktivitása fontos szerepet játszik a LINE elemek transzpozíciós mechanizmusában (6). A LINE elemek transzkripciója a SINE elemektől eltérően RNS polimeráz II által katalizált. A transzpozíció során a LINE elem RNS kópiájának 3' végét felismeri a reverz transzkriptáz majd ugyanez az enzim a cél szekvenciánál elvágva a DNS molekulát és azt primerként használva elkészíti a LINE elem DNS kópiáját. Az emlős LINE elemek nagy többsége csonka példány amelyek önálló



transzpozícióra nem képesek. Kisebbségi hányaduk teljes méretű és ezek közül néhány, mutációt okozó transzpozícióját is demonstrálni lehetett (7, 8, 9).

#### **1.2.1.1.3. Retrovírus-szerű elemek:**

Az RLE elemek megkülönböztető sajátága a repetitív elemet határoló LTR (*long terminal repeat*) szakasz. Az LTR-ek transzkripciós regulátor elemeket tartalmaznak és fontos szerepet játszanak a transzpozíció folyamatában is. Az RLE elemek egy csoportja az úgynevezett endogén retrovírusok (ERV) nagyfokú hasonlóságot mutatnak a tényleges retrovírusok szerveződéséhez. Az ERV elemek nagy valószínűséggel retrovírusokból jöttek létre a fertőző képesség elvesztése révén (10). Az emlősök endogén retrovírusai mellett az élesztőben azonosított *Ty* elem és a *Drosophila*-ból ismert *copia* elem is ebbe a csoportba sorolható. Ugyancsak az RLE-k közé tartoznak, de a retrovírusokra már kevésbé hasonlítanak, az emlősök genomjában több százezer példányban előforduló MaLR (*mammalian LTR-retroposon*) elemek (11). Az MaLR elemek és egyéb RTL-el rendelkező de már reverz transzkriptázt nem kódoló repetitív elemek mobilitása külső reverz transzkriptáz forráshoz kötött.

#### **1.2.1.2. DNS transzpozon eredetű repetitív elemek:**

Az eukarióta genom szétszórt repetitív elemeinek jelentős része DNS transzpozon eredetű. Szemben a retropozonokkal, melyek elterjedése csak az eukariótákra korlátozódik, a DNS transzpozonok már a prokariótákban is megtalálhatók. Általános jellemzőjük a transzpozon részét képező fordított végi ismétlődések (*inverted terminal repeats*). A transzpozont két oldalról határoló, az integráció célszekvenciájaként szolgáló rövid direkt ismétlődések a transzpozon típusára jellemzőek, de már nem a transzpozon integráns részei. A DNS transzpozonok transzpozíciós mechanizmusa jelentősen eltér a retropozonok mobilitásáért felelős folyamatoktól. Az úgynevezett nem replikatív transzpozíció során a transzpozon DNS kópiája kivágódik a kromoszómából és a genom más helyeire integrálódik.



A transzpozíciót egy a transzpozon által kódolt enzim az ún. transzpozáz katalizálja. A transzpozíció replikatív módon is lejátszódhat, ha a kivágódott transzpozon helyén maradó törés javítódik, templátként használva a transzpozont még tartalmazó testvér kromatidát. Prokarióta transzpozonok esetében a replikatív transzpozíció kointegrátum képződéséhez vezet. Ilyenkor a transzpozáz mellett egy, a kointegrált struktúrát feloldó enzim a rezolváz aktivitása is szükséges. Az eukarióta DNS transzpozonok legismertebb képviselői a *Drosophila* P eleme és a kukorica *Ac/Ds* rendszere. A legutóbbi időkig a rendelkezésre álló adatok arra utaltak, hogy az emlősök szétszórta repetitív elemei mind retroponon eredetűek. Újabb szekvencia analízisek azonban kimutatták, hogy a humán genomban megtalálhatók az eukarióta DNS transzpozonok két nagy osztályának, az *Ac/hobo* és a *Tc1/mariner* osztályok képviselői (12). Az azonosított elemek jelentős hasonlóságot mutatnak az adott transzpozon osztály rovarokból ismert tagjaival, ami arra utal, hogy a fajok között lejátszódó ún. horizontális terjedési mód fontos szerepet játszik ezen DNS elemek evolúciójában (13).

### 1.2.2. Tandem repetitív elemek:

A tandem repetitív elemek csoportjába eredetüket és szerkezetüket tekintve heterogén repetitív DNS szekvenciák sorolhatók. Közös sajátosságuk, hogy mindig azonosítható egy rövidebb hosszabb DNS szakasz amely tandem ismétlődik akár több százezerszer. Az ismétlődő szekvencia egység (*repeat unit*) hossza egy nukleotidtól néhány ezer tizezer nukleotidig terjedhet. Az tandem repetitív szekvenciák három fő családját különíthetjük el: (i) tandem repetitív gén-családok, (ii) szatellitek és (iii) miniszatellitek vagy "egyszerű" szekvenciák (*simple sequences*). A családokba való sorolás alapját nem szekvencia szinten megfigyelhető homológia, hanem szerveződésbeli hasonlóságok szolgáltatják. A különböző családokba sorolt repetitív elemek nemcsak az ismétlődő egység jellemzőiben, de általában kromoszómális lokalizációjukban és a genom szerveződésében betöltött szerepükben is eltérnek. A csoportok között nem állapítható meg éles átmenet. Sok esetben a besorolás

önkéntes és az ismétlődő egység mérete mellett a fentebb említett egyéb sajátosságokat is figyelembe veszik az elem osztályozásánál. Tovább bonyolítja a helyzetet, hogy adott szekvencia elemek egyes fajokban miniszatellit-szerű szerveződést mutatnak, míg más fajokban ugyanezek az elemek szatellit frakciót képeznek (14). A különböző családokba tartozó repetitív elemek hierarchikus egymásra épülését is megfigyelhetjük. Például előfordul, hogy miniszatellit blokkok részét képezik egy másik tandem repetitív elem hosszabb ismétlődő egységének.

A tandem repetitív szekvencia családok evolúcióját, jelentős különbségeik ellenére hasonló folyamatok vezérik. Az ismétlődő egységek amplifikációjához mechanizmusukat tekintve többféle folyamat vezethet. A kromoszómák testvér kromatidái között lejátszódó egyenlőtlen *crossing over* az egyik kromatidán az ismétlődő egységek számának növekedéséhez míg a másikon csökkenéséhez vezethet. Az egymást követő generációk során ez a folyamat a repetitív blokk méretének lépcsőzetes növekedését és csökkenését is eredményezheti. A repetitív egységek száma egy generáció alatt is jelentősen növekedhet úgynevezett szaltatorikus amplifikáció révén (15). Ennek az amplifikáció típusnak a pontos molekuláris mechanizmusa még nem tisztázott, de valószínűleg a repetitív blokkon belüli intrakromoszómális rekombinációt és ennek eredményeként ún. *rolling circle*-szerű struktúra kialakulását foglalhatja magában (16). A szaltatorikus amplifikáció eredményezheti a tandem repetitív szakaszokon belül megfigyelhető nagyméretű homogén blokkok kialakulását (17). Rövid repetitív egységek számának változása a DNS szintézis során, a repetitív egységek egész számú többszörösének megfelelő szálelcsúsítások révén is történhet (*polymerase slippage*) (18). A szintetizálódó szál elcsúsztatása és kihurkolódása a repetitív egységek számának növekedéséhez, míg a teplát szálé a csökkenéséhez vezethet.

Annak ellenére, hogy a fentebb részletezett amplifikációs folyamatok eredményeként a repetitív egységek száma a több ezret is elérheti, gyakran meglepő mértékű homológia figyelhető meg az egységek szekvenciái között, olyan esetekben is amikor az nem



magyarázható funkcióhoz kötött szelekciós nyomással. Ezt a jelenséget a tandem repetitív szekvenciákra jellemző ún. összehangolt evolúcióval (*concereted evolution*) értelmezhetjük (19). Az összehangolt evolúció feltételezi, hogy a nagyfokú homológia fennmaradásáért a random mutációs folyamatokkal ellentétes irányban működő homogenizációs folyamatok a felelősek. Ezen homogenizációs folyamatok molekuláris alapjául a génkonverzió és az egyenlőtlen *crossig-over* szolgálnak. A génkonverzió homogenizáló hatásához fell kell tételeznünk, hogy a nem allélikus repetitív egységek között lejátszódó szál kicserélődés során létrejövő *mismatch* javítása az egyik irányban preferált. A többszöri egyenlőtlen kicserélődés és génkonverzió a javítás során preferált szekvenciáknak a repetitív egységek populációján belüli elterjedéséhez vezethet, ugyanakkor a javítás során nem preferált mutációk rögzülését megakadályozhatja. Együttesen ezen folyamatok hatásaként a repetitív egységek szekvencia homológiája sokkal nagyobb mértékű lesz mint ami pusztán random mutációs események eredményeként létrejöhetne. Összefoglalva tehát a tandem repetitív elemek evolúcióját három tényező az amplifikáció, random mutációk és ez utóbbiakkal ellentétesen ható homogenizáció eredője határozza meg.

### **1.2.2.1 Tandem repetitív gén-családok:**

A repetitív gén-családok általánosan elterjedtek az eukariótákban. Legismertebb képviselőik a globin gének, az immunoglobulin gének, a hiszton gének, az aktin gének a rRNS és snRNS gének családjai. A gének amplifikációja fontos evolúciós következményekkel járhat. A létrejövő új kódoló szekvenciák, egy kivételével, a szelekciós nyomás alól felszabadulva szabadon változhatnak. A szám feletti kópiák a felhalmozódó mutációk eredményeként új tulajdonságokkal rendelkező fehérje molekulákat meghatározó génekké alakulhatnak. A mutációk felhalmozódása a kódoló szakaszok teljes inaktiválódásához is vezethet. Az ilyen funkcióját veszített DNS elemeket pseudo-géneknek nevezzük. Az amplifikáció eredményeként kialakuló új génkópiák általában egymás

közelében maradva csoportokat alkotnak. Egyes esetekben a kódoló szakaszok közötti régiók szerkezete és mérete változhat. Ilyen csoportos, de nem tandem elrendeződést mutatnak az emlősök globin génjei. Ezzel szemben az emlősök akrocentrikus kromoszómáinak rövid karján elhelyezkedő riboszómális RNS gén csoportok több mint 40 kbp méretű, tandem szerveződő repetitív egységekből épülnek fel. A repetitív egységek a 18S az 5.8S és 28S rRNS-eket meghatározó ~14 kbp-nyi kódoló szakaszból és ~30 kbp nem kódoló ún. összekötő szakaszból állnak. Szekvencia analízisek kimutatták, hogy a repetitív egységeknek nemcsak a kódoló szakaszai hanem az összekötő régiói is gyakorlatilag azonosak. Ezt a rendkívül nagyfokú szekvencia homogenitást a tandem repetitív DNS elemekre jellemző összehangolt evolúciós folyamatokkal magyarázhatjuk. A rRNS génekhez hasonló, tandem szerveződésű gén csoportokat alkotnak még például a hiszton gének és az U2 snRNS gének (20).

#### **1.2.2.2. Szatellitek:**

Az eukarióta DNS CsCl<sub>2</sub> vagy CsSO<sub>4</sub> sűrűség gradiens centrifugálás során általában több különböző sűrűségű frakcióra válik szét. A genom DNS tartalmának nagy részét tartalmazó fő frakció mellett megjelenő kisebb frakciókat szatelliteknek nevezzük (*classical satellites*). A klasszikus szatelliteket alkotó DNS frakciók sűrűsége eltér a genomikus DNS átlagos sűrűségétől. A szatellitek szekvencia analízise kimutatta, hogy ezek a frakciók olyan DNS fragmenteket tartalmaznak amelyek rövid szekvencia részletek nagyszámú tandem ismétlődéséből épülnek fel. Ez a szerveződési mód az átlagostól eltérő bázis összetételt és a genomikus DNS átlagos sűrűségétől eltérő, az adott szatellit frakcióra jellemző sűrűség kialakulását eredményezi. A szatellit DNS frakciók mennyisége fajonként változik és 5-10%-tól egészen 50%-ig terjedhet. Ha a sűrűség gradiens centrifugálás során a gradienst alkotó pufferhez interkalálódó DNS festékeket (pl. etidium bromid, Hoechst 33258), antibiotikumokat (pl. aktinomycin), vagy nehéz fém ionokat (pl. ezüst, higany) adunk az



elválasztható szatellit frakciók száma megemelkedik. Az ilyen körülmények mellett megjelenő új szatellit frakciókat rejtett szatelliteknek (*cryptic satellites*) nevezzük. A tandem szerveződést mutató, szatellit-szerű repetitív DNS frakciók egy része a fentebb ismertetett feltételek mellett sem választhatók el a genomikus DNS fő frakciójától sűrűség gradiens centrifugálás segítségével. Ezért az eredeti definíciót módosítva, a hasonló szerveződési módra helyezve a hangsúlyt, szatellitnek tekinthetünk minden, rövidebb hosszabb ismétlődő alapegységgel jellemezhető tandem repetitív DNS elemet, függetlenül attól, hogy sűrűség gradiens centrifugálás során elválik-e a genomikus DNS fő frakciójától, vagy sem.

A szatellit szekvenciák szerveződése legegyszerűbben restriktív enzimekkel tanulmányozható. A szatellit frakciók részleges restriktív emésztést követő gélelektroforézis a DNS fragmentek létraszerű mintázatát mutatja ki. Az egyes fragmentek mérete a repetitív egység méretének egészszámú többszöröse. A létraszerű mintázat sok esetben teljes emésztés után is megmarad. Ennek a magyarázata az, hogy a repetitív egységek habár nagyfokú hasonlóságot mutatnak, nem teljesen azonosak. Mutációk elronthatják a használt restriktív enzim felismerő helyeit egyes repetitív egységekben. Ilyenkor a kérdéses enzimmel való teljes emésztés, a monomer egynél nagyobb számú többszörösének megfelelő méretű fragmenteket is eredményez. Ha több, különböző felismerő helyű restriktív enzimmel történő részleges vagy teljes emésztést követően jellegzetes, létraszerű mintázatot látunk, az az adott repetitív szekvencia elem tandem szerveződésének fontos bizonyítéka.

Klónozott szatellit DNS elemekkel végrehajtott *in situ* hibridizációk kimutatták, hogy a szatellitek túlnyomórészt a kromoszómák pericentrikus, kisebb részben szubtelomerikus régiókban lokalizálódnak (21). Ezek a jellegzetes festődést mutató ún. konstitutív heterokromatikus kromoszómaszakaszok kevés gént tartalmaznak és a sejtciklus S fázisának végén replikálódnak. A heterokromatin általában negatív hatással van a gének expressziójára (*position effect variegation: PEV*). Ezt a *Drosophilában* régóta ismert jelenséget (22) újabban kimutatták hasadó élesztőben (23) is és bizonyos adatok arra utalnak, hogy néhány humán

öröklődő betegség kialakulásában is szerepe lehet PEV-szerű jelenségnek (24). A PEV mechanizmusára szolgálhatnak adatokat azok a kísérletek amelyekben megállapították, hogy transzének tandem elrendeződése heterokromatin-szerű konformáció kialakulását és a transzén csökkent expresszióját eredményezheti (25, 26). Habár általánosan a heterokromatin és a genetikai inaktivitás fogalma összekapcsolódik újabban számos jelentős kivételt találtak ez alól a szabály alól. *Drosophilában* azonosítottak géneket amelyek megfelelő szintű expressziójához elengedhetetlen heterokromatikus lokalizációjuk (27). Emlősökben is kimutatható, hogy a pericentrikus heterokromatin közelében elhelyezkedő gének stabil expressziót mutathatnak (28).

A szatellitek jellegzetes lokalizációja és lehetséges funkciói közötti kapcsolat régóta képezi vizsgálatok tárgyát. Egyre több adat halmozódik fel, amelyek a kromoszóma-szegregáció és a pericentrikus szatellitek kapcsolatára utalnak (29) (lásd később).

### **1.2.2.3. Miniszatellitek ("egyszerű" szekvenciák):**

A miniszatellitek és a szatellitek közötti határ a repetitív egységek méretét tekintve nehezen húzható meg. Habár általánosságban elmondhatjuk, hogy a miniszatellitek esetében az ismétlődő egység mérete a tíz bázispárt nem haladja meg, ismertek klasszikus szatellitek, amelyek fő komponense ugyancsak tíz bázispárnál rövidebb. A repetitív egység méreténél sokkal megbízhatóbb elhatárolást tesz lehetővé ha a repetitív egységek számát és a repetitív blokkok kromoszómális lokalizációját vizsgáljuk meg. Míg a szatellitek koncentráltan, nagy heterokromatikus blokkokba tömörülve, általában a pericentrikus régiókban helyezkednek el, addig a miniszatellitek sokszor az eukromatinban szétszórva, rövid blokkokat alkotva fordulnak elő (30).

A miniszatellitek vagy "egyszerű" szekvenciák legelterjedtebb képviselője a  $(TG)_n$  szekvencia elem. Kimutatták, hogy ez az eddig vizsgált összes eukariótában előforduló repetitív elem fiziológiás sókoncentráció mellett balmenetes Z-DNS struktúra felvételére



képes (31). Ugyancsak sikerült fehérjéket azonosítani, amelyek specifikusan kötődnek ehhez az alternáló purin-pirimidin szakaszhoz (32). Feltételezések szerint ez a szekvencia elem résztvehet gének transzkripció szabályozásában, míg más adatok a DNS szintézis iniciációjában illetve a kromoszómális rekombinációban betöltött potenciális szerepét hangsúlyozzák. Újabban különböző trinukleotid szekvencia elemek kerültek az érdeklődés középpontjába. Több humán öröklődő betegségről valamint rákos transzformációhoz vezető mutációkról derítették ki, hogy a defektusok tandem trinukleotid egységeket tartalmazó szakaszok instabilitásaira vezethetők vissza (33). Az eukarióták többségének kromoszómái  $(TTAGGG)_n$  hexanukleotid egységekből álló szakaszokban végződnek (34). Funkcionális tesztek bizonyították, hogy ennek a repetitív elemnek 250-1000 kópiájából álló terminális DNS szakaszok a telomer szerepét látják el számos fajban (35) (lásd később). Habár a miniszatellitek többségének szerepéről nagyon keveset tudunk, nagyfokú polimorfitásuk révén újabban széleskörű alkalmazást nyertek genetikai térképezésekben és igazságügyi orvostani vizsgálatokban (36).

### ***1.3. Repetitív szekvenciák és a kromoszómaszerkezet kapcsolata:***

A tankönyvi definíció szerint a kromoszóma egy adott organizmus genetikai információjának egészét vagy egy részét hordozó nukleoprotein komplex. Közelebbről megvizsgálva kiderül, hogy a kromoszóma több mint egy élőlény génjeinek puszta gyűjteménye. A kromoszómák több funkciót ellátó sejtorganellumok. Legrégebben felismert és egyben egyik legfontosabb szerepük a genetikai információnak az utód sejtek közötti pontos eloszlásának biztosítása a sejt osztódása során. A kromoszómák nem passzív résztvevői az osztódásnak. Mind a kromoszómális mozgásokért felelős mechanizmusoknak, mind a folyamat szabályozásának aktív, központi szereplői (37). Újabban a kromoszómáknak szerepet tulajdonítanak az interfázisos sejtmagban megfigyelhető ún. sejtmagi rend kialakulásában is (38). Jelenlegi ismereteink szerint az eukarióta kromoszómák

nélkülözhetelen struktúrális és funkcionális alkotóelemei a centromer, a telomer és a DNS szintézis kezdőpontjaiként szolgáló szekvencia elemek (ARS: *autonomous replicating sequences*) (39, 40). A következőkben ezen három fő kromoszómális alkotóelem és a korábban ismertett repetitív szekvencia elemek kapcsolatát kívánom részletesebben ismertetni.

### 1.3.1. DNS replikációs origó:

A DNS replikáció kezdőpontjait az elmúlt évek kutatásai eredményeként prokariótákban és az alacsonyabbrendű eukarióták közül az élesztőben ma már molekuláris szinten tudjuk definiálni (41). Habár az utóbbi években jelentős előrehaladás történt az emlős vírusok DNS szintézisét kezdeményező starthelyek jellemzésében (42), a magasabbrendű eukarióták kromoszómális replikációs origóiról még mindig viszonylag keveset tudunk. A legintenzívebben tanulmányozott emlős replikációs origó a kínai hörcsög dihidrofolát reduktáz (DHFR) génjének közelében helyezkedik el. Ennek a kromoszóma szakasznak az amplifikációja gyakran figyelhető meg hörcsög sejtek metotrexát szelekciója során (43). Két-dimenziós gélelektroforézissel kimutatható, hogy a DNS szintézis egy több tíz kilobázispár hosszúságú DNS szakaszon belül több helyen indul el (44). Nagyobb felbontású kísérleti módszerekkel a legkorábban replikálódó DNS szakaszok mérete 3-6 kb-nak adódott. Az így módon azonosított DNS fragmentek szekvencia analízise különböző repetitív elemek koncentrációját mutatta ki a replikációs start helyek közelében. A transzpozonszerű elemek (ORR1) mellett, *Alu* elemekkel rokon szekvenciák és Z-DNS szerkezet valamint DNS triplexek kialakítására képes "egyszerű" szekvencia elemek voltak azonosíthatók (45). Összefoglalva, habár az emlősökben eddig még nem sikerült egyértelműen azonosítani az élesztő ARS szakaszaihoz hasonló, definiált replikációs origókat, azt már az eddigi adatok is erőteljesen sugalmazzák, hogy különböző repetitív elemek, különösen a B-DNS struktúrától



eltérő szerkezetek kialakítására képes "egyszerű" szekvenciák fontos szerepet játszhatnak az emlős DNS szintézis iniciációjában.

### 1.3.2. Telomer:

A DNS szintézisét katalizáló polimeráz működésének elengedhetetlen feltétele a mintaként szolgáló templát DNS szál és a templáttal komplementer, szabad 3' hidroxil csoportban végződő primer szál jelenléte. *In vivo* DNS szintézis során RNS polimeráz által szintetizált rövid RNS szakaszok szolgálnak a DNS polimeráz számára primerként. Ezek az RNS szakaszok később eltávolítódnak a DNS-ből és a helyükön maradó réseket DNS polimeráz tölti ki. A DNS polimeráz a DNS szintézisét 5'→3' irányban katalizálja. A kettősszálú DNS replikációja során a 3'→5' polaritású szálon a komplementer szál szintézise 5'→3' irányban folyamatosan történhet (*leading strand*). A másik templát szálon a DNS szintézise 3'→5' irányban nem folyamatos (*lagging strand*). Ez a DNS szál rövid szakaszok ún. Okazaki fragmentek formájában szintetizálódik. Minden Okazaki fragment egy RNS primerrel indul amit a DNS polimeráz 5'→3' irányban meghosszabbít. A DNS polimeráz fentebb ismertetett katalitikus sajátságai folytán lineáris DNS molekulák 5' végeinek replikációja nem lehet teljes. Az egymást követő replikációs ciklusok során ez a DNS molekulák méretének fokozatos csökkenéséhez vezet (46).

Az evolúció során az eukarióták lineáris kromoszómáinak végén az inkomplett DNS replikáció hatását kivédő speciális struktúra, a telomer alakult ki. Jelenlegi ismereteink szerint az eukariótákban három fő telomer típus figyelhető meg (47, 48). Az esetek többségében a telomer rövid, 5-8 bp méretű, asszimmetrikus G megoszlást mutató, tandem repetív egységekből épül fel. Ez a telomer típus jellemző az emlősökre, a csillós egysejtűekre, a magasabbrendű virágos növényekre, a rovarok és élesztők egy részére. Az emlősökben a kromoszómák végén a TTAGGG hexanukleotid több száz tandem kópiája képezi a telomert (34). Kimutatható, hogy a kromoszómák 3' hidroxil csoportot hordozó

egyszálú DNS szakaszban végződnek (49). Egyes adatok arra utalnak, hogy a telomerikus repetitív egységekhez hasonló G gazdag, egyszálú DNS szakaszok nem Watson-Crick bázis párosodások révén négyzálú ún. G-kvartet struktúra kialakítására képesek (50). Ma még nem tisztázott, hogy ez a struktúra *in vivo* megtalálható-e a kromoszómák végén. A kromoszómák replikációja során folyamatosan elvesző telomerikus DNS szakaszokat egy speciális reverz transzkriptáz, a telomeráz pótolja (51). A telomeráz integráns részét képezi egy RNS molekula amely templátként szolgál a telomerikus repetitív egységek szintéziséhez. Emlősökben telomeráz aktivitás csak az ivarsejtekben mutatható ki. A telomeráz aktivitás hiánya a testi sejtekben a kromoszómák telomerjeinek folyamatos rövidüléséhez vezet. Egyes elméletek szerint a telomer méretének csökkenése összefüggésben van a sejtek normális öregedési folyamataival (52). Újabban kimutatták, hogy a korlátlan osztódási kapacitással rendelkező transzformált rákos sejtek a normális testi sejtektől eltérően telomeráz aktivitást mutatnak (53). Feltételezik, hogy a sejt telomeráz aktivitásának visszanyerése a rosszindulatú transzformáció fontos lépését képezi.

Az elvesző telomerikus DNS pótlásának telomeráztól független módját írták le *Saccharomyces cerevisiae*-ben. Habár feltételezhetően ez az élesztő faj is tartalmaz telomerázt, kimutatható, hogy a kromoszómák szubtelomerikus régióiban előforduló transzpozon-szerű ún. Y' elemek között lejátszódó, RAD52-függő rekombinációs folyamatok is hozzájárulnak a telomerikus DNS szakaszok meghosszabbodásához (54).

Egyes egysejtű gombákban (*Candida*, *Kluyveromyces*) a telomerikus repetitív egységek mérete az előbb ismertetettnél hosszabb (8-25 bp) és nem figyelhető meg a G nukleotidok szálak közötti eloszlásának erős asszimetriája. A hosszabb repetitív egység ellenére a telomer szintézise az előzőekben bemutatottakhoz hasonlóan valószínűleg telomeráz által katalizált.

A harmadik telomer típus a *Drosophilára* és valószínűleg más *Diptera* fajokra jellemző. Feltételezések szerint ezen fajok evolúciójuk során elvesztették a rövid repetitív



egységekből felépülő, telomeráz által fenntartott telomer szerkezetet és helyette egy mobilis repetitív elemekre alapuló, kromoszóma végeket védő mechanizmusra tettek szert. A kromoszómák DNS vesztését időről-időre a kromoszómák végeire ugró retroponok ellensúlyozzák. Eddig két speciális retropon típusot azonosítottak amelyek ezt a feladatot ellátják. Mindkét elem az emlősök LINE elemeihez hasonló LTR-rel nem rendelkező retropon. Az elsőként leírt HeT-A elem nem kódol reverz transzkriptázt emiatt mobilitása külső enzim forráshoz kötött (55). Az újabban azonosított TART elem saját reverz transzkriptáz kódolására képes autonóm retropon (56).

Az ismertetett telomer típusok különbségeik ellenére figyelemre méltó közös sajátysága, hogy mindegyik esetben reverz transzkripcióra épülő mechanizmus biztosítja a kromoszóma végeket védő struktúra fennmaradását.

### 1.3.3. Centromer:

Az eukarióta kromoszómák centromerikus régiója központi szerepet tölt be a sejtosztódás során lejátszódó kromoszómális mozgások végrehajtásában és szabályozásában egyaránt. Genetikai értelemben a centromert egy olyan *cis* hatású elemként definiálhatjuk, amely biztosítja a kromoszómák mitózis és meiózis során lezajló pontos szegregációját. A magasabbrendű eukarióták esetében a centromer citológiai szinten is definiálható (57). A fajok jelentős részében a centromer a metafázisos kromoszómák ún. elsődleges befűződésének közelében helyezkedik el. Elektronmikroszkóppal az elsődleges befűződés két oldalán három rétegből felépülő korong-szerű struktúra, a kinetochor figyelhető meg. Ez az összetett felépítésű, nagyrészt fehérjékből álló szerkezet és az elsődleges befűződés közelében elhelyezkedő centromerikus heterokromatin, szerkezeti és funkcionális egységet képeznek, amit centromer-kinetochor komplexnek nevezünk. A centromer-kinetochor komplex vázlatos szerkezetét az első ábra mutatja.

### 1.3.3.1. A centromer-kinetochor komplex funkciói:

A kinetochor sejtosztódásban betöltött szerepe többrétű. A sejtosztódás során kialakuló magorsó, mikrotubulosokból álló húzófonalai közvetlenül ehhez a struktúrához kapcsolódnak. A kinetochor azonban nem csak passzív kapcsolószerkezet funkcióját tölti be, hanem aktívan is részt vesz a kromoszómális mozgásokért felelős erőhatások létrejöttében. Ezt az ún. mechano-kémiai motor funkciót az ATP-függő motor fehérjével, a kinezinnel nagyfokú homológiát mutató kinetochor fehérjék léte is alátámasztani látszik (58).

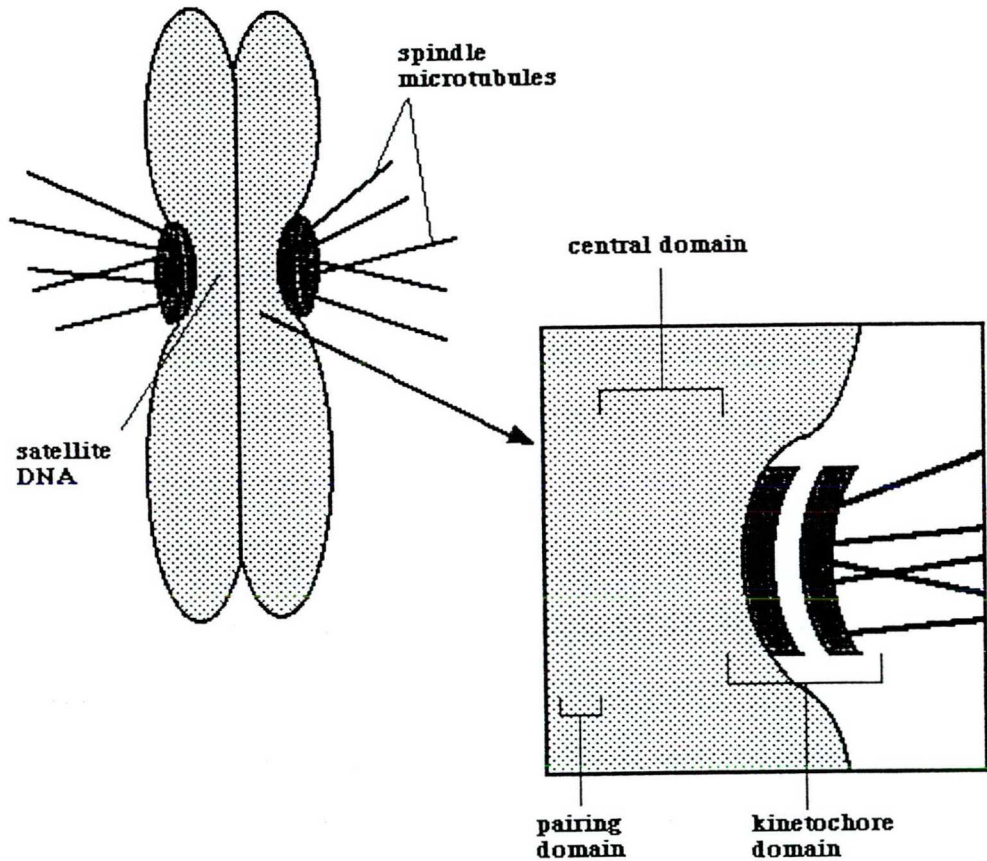
A kromoszómák pontos szegregációjának biztosításában fontos szerepe van a metafázis-anafázis határán működő sejtciklus ellenőrző pontnak. Ha a metafázis végén a sejtben még vannak a húzófonalához nem megfelelően kapcsolódott kromoszómák a sejt osztódása a metafázis-anafázis határán leáll. A sejtosztódás mindaddig nem folytatódik amíg az összes kromoszóma megfelelően nem kapcsolódik a húzófonalához. Feltételezik, hogy a kromoszómák kapcsolódását ellenőrző folyamatban a centromer-kinetochor komplex bizonyos komponensei központi szerepet töltenek be. Újabb eredmények arra utalnak, hogy a jelző mechanizmus bizonyos kinetochor fehérjék foszforilációs állapotának változásait foglalja magában (59).

A kromoszómát alkotó két testvérkromatida a sejtosztódás során centromerikus régiókban maradnak legtovább egymáshoz kapcsolva. A kromatidák egymással érintkező belső oldalán speciális, feltételezhetően kapcsoló funkciót ellátó fehérjék mutathatók ki (60). Ezek a fehérjék valószínűleg végrehajtó szerepet töltenek be a kromatidák megfelelő időben történő elválását szabályozó folyamatban. A kromoszómák számtartó, illetve redukciós szegregációját tehát alapvetően a centromerikus régióban lezajló időzített kromatida elválás határozza meg.

A centromerikus heterokromatin interfázisban megőrzi kompakt állapotát. Centromerikus fehérjéket felismerő ellenanyagok a metafázisos kromoszómák centromerjeihez hasonló, kettős, pontszerű struktúrákat ismernek fel interfázisban is. További



citológiai adatok azt is mutatják, hogy adott kromoszómák centromerikus régióinak egymáshoz viszonyított elhelyezkedése az interfázisos sejtmagban nem random (61). Ezek a megfigyelések együttesen arra utalnak, hogy a centromernek szerepe lehet egyfajta sejtmagi rend kialakulásában. Egyes feltételezések szerint ennek a sejtmagi kompartmentalizációnak a szervezésében a centromer közelében elhelyezkedő repetitív szekvenciáknak döntő szerep jut (38).



1. ábra. A magasabb rendű eukarióták centromer-kinetochor komplexének vázlatos felépítése.

### 1.3.3.2. A centromer-kinetochor komplex molekuláris felépítése:

A centromer-kinetochor komplex molekuláris összetételének vizsgálatában az utóbbi években jelentős előrehaladás történt. Ezek a kutatások ma már elegendő információt szolgáltatnak ahhoz, hogy több élesztő fajban a centromert molekuláris szinten is definiálni tudjuk. Habár a magasabbrendű eukarióták esetében is jeletős eredmények születtek, különösen a kinetochor fehérje komponenseinek azonosítása terén (62, 63), DNS szinten még mindig távol állunk attól, hogy a centromert olyan pontossággal definiálhassuk mint amit az élesztő esetében már megtehetünk. A következőkben elsősorban a különböző eukariótákban azonosított centromerikus DNS elemeket ismertetem. A centromer-kinetochor komplex eddig azonosított fehérjeinek tulajdonságait és feltételezett funkcióit az 1. táblázat foglalja össze.

Mammalian centromere proteins			
Protein	Size(kDa)	Localization	Comments
CENP-A	17	Centromere	Histonelike protein
CENP-B	80	Central domain	Alpha satellite DNA-binding protein
CENP-C	140	Inner kinetochore plate	Specific to active centromeres; DNA-binding protein
CENP-D	50	Kinetochore	Regulator of chromatin condensation?
CENP-E	312	Outer kinetochore plate	Kinetochore motor protein
CENP-F	400	Outer kinetochore plate	Initiator of centromere and kinetochore assembly?
INCENPs	96, 100	Pairing domain/inner chromatids	Sister chromatid association; cytoskeleton assembly
CLiPs	?	Pairing domain	Sister chromatid linking

**1. táblázat. Az ismert emlős centromer-kinetochor fehérjék biokémiai sajátosságai, funkciói és citológiai lokalizációja.**

Centromerikus DNS szekvenciákat először a közöséges sör élesztőben (*Saccharomyces cerevisiae*) azonosítottak azon képességük alapján, hogy cirkuláris plazmid

molekulákba klónozva őket lehetővé tették azok pontos mitotikus szegregációját (64). A még teljes funkciót mutató élesztő centromerikus DNS minimális mérete 155-160 bp (65). *In vivo* a centromerikus DNS speciális fehérjékkel nukleáz rezistens komplexet képez, amelyet két oldalról rögzített helyzetű nukleosómák határolnak. Az élesztő centromer három konzervált szekvenciái elemből épül fel: CDEI, CDEII és CDEIII. A CDEI nyolc bázispár hosszú szekvencia elem. Deléciója a kromoszóma mitotikus stabilitását ~ 60-ad részére csökkenti, míg a meiótikus stabilitást csak kis mértékben befolyásolja. A CDEII elem pontos nukleotid szekvenciája kevésbé konzervált, viszont mérete és bázisösszetétele fontos. A pontos kromoszóma szegregáció biztosításához az elem méretének 78-86 bp-nak, az A-T tartalomnak legalább 90% kell lennie. A CDEII elem jelenléte meiótikus osztódás szempontjából esszenciális. A magas A-T tartalom feltételezések szerint a CDEII görbülését váltja ki, ami viszont a másik két centromerikus szekvencia elemhez kapcsolódó egyéb faktorok megfelelő kölcsönhatásához lehet szükséges. A CDEII-t követő CDEIII 26 bp hosszú, szekvenciájában tükörszimmetriát mutató DNS elem. A szimmetria középpontját képező C nukleotid mutációja a centromer funkció teljes elvesztésével jár.

A hasadó élesztő (*Schizosaccharomyces pombe*) kromoszómái hasonlóan a magasabbrendű eukarióták kromoszómáihoz sejtosztódás során kondenzálódnak. Egyes feltételezések szerint ez a sajátosság adja a magyarázatát annak, hogy ez az élesztőfaj a magasabbrendű eukarióták felé átmenetet képező, komplexebb centromer szerveződést mutat. A centromerikus DNS mérete 40 kb és 100 kb között változik kromoszómától függően (66). A eltérő méret ellenére a hasadó élesztő mindhárom kromoszómájának centromerje hasonló módon szerveződik. A 4-7 kb méretű egyedi központi szekvenciát mindkét oldalról, egymáshoz viszonyítva ellentétes orientációjú repetitív elemek veszik körül. Deléciós kísérletek bizonyították, hogy a központi szekvencia egy része és az ún. K-típusú repetitív elem egy 2.1 kb-os fragmentje együttesen képesek ellátni a centromer funkcióját (67). A hasadó élesztő centromerje is különleges kromatin szerkezetet mutat. A kromatin szerkezet és



a centromer funkció kapcsolatát húzza alá az a megfigyelés, hogy centromerikus DNS szakaszokat tartalmazó plazmidok mitotikus stabilitását a centromer körül kialakuló kromatin szerkezet nagy mértékben befolyásolja. A kromatin szerkezet változása akár a centromer teljes inaktiválódását is eredményezheti. Az aktív kromatin konformáció kialakulását a K-típusú repetitív elem *enhancer-szerű* módon segíti elő (68). Egyes feltételezések szerint a repetitív elemek ugyanakkor szerepet játszanak a testvér kromatidák összetartásában is (65).

A *Drosophila* centromerikus régiója már a többsejtű eukariótákra jellemző komplex szerveződést mutatja. Az esszenciális centromerikus DNS szakasz azonosításához nagymértékben hozzájárult, egy ~1.3 Mb méretű *Drosophila* marker kromoszóma genetikai analízise (69). Ezen kromoszóma különböző deléciós származékai eltérő mitotikus és meiótikus stabilitást mutatnak. A deléciós származékok molekuláris vizsgálata vezetett annak a ~420 kb méretű kromoszóma régióhoz az azonosításához, amely képes a teljes centromer funkció ellátására. A minimális centromer funkció ezen szakaszon belül, egy még nem teljesen definiált ún. komplex DNS-ből (egyedi és közepesen repetitív szekvenciák keveréke) felépülő, ~220 kb méretű szakaszhoz rendelhető. Feltételezik, hogy ez a régió felelős a kinetochor kialakításáért. Az intakt kromoszómákra jellemző mitotikus és meiótikus stabilitás eléréséhez további ~200 kb-nyi szatellit DNS jelenléte is szükséges. A szatellit DNS legvalószínűbb funkciója a testvér kromatidák közötti kapcsolat kialakítása.

Az emlősök kromoszómaiknak pericentrikus régiójában heterogén, tandem szerveződést mutató, szatellit szekvencia családok koncentrálnak (70, 71) (2. táblázat). A funkcionális centromer azonosítását mindeddig nagyban megnehezítette, hogy a kiterjedt kutatások ellenére sem sikerült az emlősök többségében konzervált centromerikus DNS szekvenciákat azonosítani. A konzervált centromerikus DNS elemek hiánya látszólagos ellentmondásban van azzal a ténnyel, hogy evolúciósan viszonylag távol álló emlős fajok között is könnyen létrehozhatók sejtibridek, amelyekben mindkét faj kromoszómái stabilan képesek fennmaradni. Ezt a ellentmondást látszanak feloldani azok az újabb kutatások

amelyek ha hosszabb, általánosan előforduló DNS elemeket nem is rövid, potenciális fehérje-kötő helyekként szolgáló, konzervált oligonukleotidok előfordulását demonstrálták egymástól távol álló eukarióták centromerikus DNS frakcióiban (72, 73).

Human Centromeric Satellite DNA Families		
Satellite	Repeat unit	Chromosomal location
Satellite I	42 bp	Acrocentric chromosomes, chromosomes 3, 4
Satellite II-III	5 bp (GGAAT)	All chromosomes
$\alpha$ -satellite	171 bp	All chromosomes
$\beta$ -satellite	68 bp	Acrocentric chromosomes, chromosomes 1, 3, 9, Y
$\gamma$ -satellite	220 bp	Chromosomes 8, X
48 bp satellite	48 bp	Chromosomes 2, 13, 14, 15, 21, 22, Y
Sn5 satellite	N/D	Acrocentric chromosomes, chromosomes 2, 20

Mouse Centromeric Satellite DNA Families		
Satellite	Repeat unit	Chromosomal location
Major satellite	234 bp	All chromosomes
Minor satellite	120 bp	All chromosomes

Hamster Centromeric Satellite DNA Families		
Satellite	Repeat unit	Chromosomal location
Hexamer satellite	6 bp (TTAGGG)	All chromosomes except chromosomes 1 and 2
HC2 satellite	2.8 kbp	Chromosome 2

2. táblázat. Emlős centromerikus szatellit DNS családok ismétlődő egységeinek mérete és kromoszómális lokalizációja

A legintenzívebben tanulmányozott emlős centromerikus DNS a főemlősök kromoszómáinak pericentrikus régióiban megtalálható  $\alpha$ -szatellit. A ~170 bp hosszú monomerre épülő szekvencia család kromoszóma specifikus magasabbrendű repetitív egységekből (*higher order repeat unit*) álló, rövidebb-hosszabb összefüggő blokkok formájában jelen van az összes humán kromoszómán (74). A monomerek egy részében



azonosítható egy 17 bp méretű DNS elem amely specifikusan köti az emlősök egyik konzervált centromer fehérjéjét (CENP-B box) (75, 76). Habár ez a szekvencia elem előfordul egyéb emlősök centromerikus szatellit szekvenciáiban is (77, 78) látszólag teljesen hiányzik a humán Y kromoszómáról (79). A CENP-B box konszenzus szekvenciájának megfelelő oligonukleotiddal végrehajtott hibridizációs kísérletek ugyancsak a CENP-B box hiányára utaltak kínai hörcsögben (80). A CENP-B fehérje és transzpozázok között újabban kimutatott szekvencia homológia a CENP-B fehérje egyéb, a centromer funkcióval közvetlenül nem kapcsolódó szerepét valószínűsíti (81). Az  $\alpha$ -szatellitnek a centromer funkciójában betöltött esszenciális szerepét más kísérleti eredmények is megkérdőjelezik. Azonosítottak olyan természetesen előforduló humán marker kromoszómákat amelyeken  $\alpha$ -szatellit nem detektálható (82, 83). Meg kell említenünk ugyanakkor azokat a kísérleteket amelyekben kimutatták, hogy emlős kromoszómákra integrálódó  $\alpha$ -szatellit a centromerre jellemző bizonyos morfológiai sajátosságokat mutatja (84, 85, 86). Összegezve ezek az adatok valószínűve teszik, hogy habár az  $\alpha$ -szatellit hozzájárulhat a centromer bizonyos sajátosságainak a kialakításához más, ma még esetleg nem azonosított szekvencia elemek is szükségesek a funkcionális emlős centromer felépítéséhez.

## 2. CÉLKITŰZÉSEK

A funkcionális emlős centromer azonosítását, új centromerikus DNS szakaszok izolálását nagyban megnehezíti a szatellit szekvenciák pericentrikus lokalizációja. A repetitív elemek, különösen a szatellitek hagyományos vektor rendszerekben klónozva általában nagymértékű instabilitást mutatnak. Mindez a centromerikus DNS elemek alul-reprezentáltságához vezet az ilyen vektorok felhasználásával készült génbankokban. Ezt a



problémát enyhítheti ha centromerikus DNS elemekben dúsított kromatin frakciót használunk parciális DNS bankok létrehozásához.

Az itt ismertetett kísérletek során egy új módszert alkalmaztunk centromerikus DNS elemekben dúsított kromatin frakció izolálására kínai hörcsög sejtekből. A módszer segítségével azonosítottunk egy új kromoszóma specifikus centromerikus DNS elemet, majd elvégeztük annak molekuláris jellemzését.

### **3. ANYAGOK ÉS MÓDSZEREK**

#### **3.1. Emlős sejtek tenyésztése:**

A CHO-K20, HeLa, LMTK<sup>-</sup> sejtek tenyésztése 10% FCS-t (*Fetal Calf Serum*) tartalmazó F12 tápoldatban történt. A concanavalin A (5-10mg/ml) indukált kínai hörcsög lépsejteket RPMI tápoldatban 72 órán át növesztettük.

#### **3.2. Kínai hörcsög kromoszómák *in vivo* fragmentációja, centromerikus kromoszóma darabok immunoprecipitációja:**

$3-4 \times 10^8$  CHO-K20 sejtet 20 órán át 2mM hidroxürea jelenlétében növesztettünk, majd a tápoldatot 5 mM koffeinnel és 0.5  $\mu$ g/ml kolchicinnel egészítettük ki. 5.5 óra után a mitotikus sejteket centrifugálással összegyűjtöttük (10', 200g). A sejteket glicin pufferben (10mM glicin, 1% Hexilén-glikol, pH:8.4-8.6) egyszer mostuk, majd ugyan ebben a pufferben 1mM PMSF (fenil-metil-szulfonil-fluorid) jelenlétében 10'-ig 37<sup>0</sup>C-on inkubáltuk. Ezt követően 0.1% Triton-X 100 detergenst adtunk a sejtekhez, majd a szuszpenziót fecskendővel 23G méretű tűn többször átnyomtuk. A sejtek feltáródását mikroszkóppal ellenőriztük. A sejtmagokat, sejttörmeléket és intakt kromoszómákat centrifugálással távolítottuk el (1500g, 20'). A felülúszót Protein-A Sepharose CL-4B (*Pharmacia*) gyantával

kapcsolt LU 851 humán anti-centromer szérummal (87) kevertük össze és szobahőmérsékleten 3 órán át inkubáltuk gyenge rázatással. Az inkubációt követően a gyantát Tween-PBS-el (137mM NaCl, 2.7mM KCl, 8.1mM Na<sub>2</sub>HPO<sub>4</sub>, 1.67mM NaH<sub>2</sub>PO<sub>4</sub>, 0.05% Tween20) 3-szor mostuk (2000 rpm, 5'). A mosás után a pellethez 0.1% SDS-t és 100µg/ml Proteináz-K-t tartalmazó TE puffert (88) adtunk és 14-18 órán át 50<sup>0</sup>C-on inkubáltuk. Fenol extrakciót követően a DNS-t 0.2M NH<sub>4</sub>-acetát jelenlétében 2.5tf 96%-os etanollal kicsaptuk. A kapott DNS-t NaeI restriktációs enzimmal történő emésztés, majd Klenow polimerázzal való kezelést követően pUC19 plazmid SmaI helyére klónoztuk.

### 3.3. *In situ* hibridizáció:

**Metafázisos kromoszóma preparátum készítése:** Kolchicin kezelést követően a tenyésztő csésze aljáról leváló mitotikus sejteket centrifugálással összegyűjtöttük (200g, 10'). A sejteket 75mM KCl oldatban 15'-ig szobahőn hipotonizáltuk majd centrifugálás után (200g, 10') -20<sup>0</sup>C-ra lehűtött metanol/ecetsav (3:1) oldatban szuszpendáltuk és 15-20'-ig jégen inkubáltuk. A sejteket centrifugálással összegyűjtöttük (200g, 10') és 2-3ml fixálóban szuszpendáltuk. A sejtuszuspenziót tiszta tárgylemezre cseppentettük majd a preparátumokat sűrített levegővel szárítottuk.

**Hibridizációs próba jelölése:** A próbaként használt DNS-be (0.5-1µg) nick-transzlációval biotin-11-dUTP-t építettünk. A jelölési reakcióhoz Amersham nick-transzlációs kitet használtunk. A reakciót <sup>3</sup>H-dGTP beépülésének mérésével követtük. A reakció EDTA-val történő leállítás után a beépületlen nukleotidokat Sephadex G-50 oszlopon választottuk el a DNS fragmentektől. A megfelelő frakcióból a jelölt DNS fragmentet etanollal kicsaptuk majd 10ng/ml koncentrációban TE pufferben oldottuk.

**Hibridizáció:** A metafázisos kromoszóma preparátumokat 100µg/ml RN-az A-t tartalmazó 2×SSC-ben (88) 37<sup>0</sup>C-on 1 órán át inkubáltuk, növekvő koncentrációjú etanol sorozattal dehidráltuk (70-90-100% etanol, 5-5') majd vákumban szárítottuk. Ezt követően a

preparátumokat 70-75°C-on 2'-ig denaturáló oldatban (70% formamid, 30% 2×SSC) inkubáltuk. A denaturáló oldatból a lemezeket 0°C-os 70%-os etanolba helyeztük majd az előbb említett módon dehidráltuk. Szárítás után a preparátumokra helyeztük az előzetesen 75°C-on 5'-ig denaturált próba elegyet (50% formamid, 10% dextrán-szulfát, 2×SSCB {10×SSCB: 1.5M NaCl, 0.15M Na<sub>3</sub>-citrát, 0.25M Tris-HCl pH:7.4, 0.5mM EDTA}, 0.5mg/ml tRNS, 3ng/ml biotinilált próba DNS; 40μl próba elegy/lemez), fedőlemezzel lefedtük őket és végül a fedőlemezek peremeit gumi ragasztóval lezártuk. A lemezeket 37°C-on nedveskamrában inkubáltuk 4-16 órán át.

**Immunodetektálás:** A fedőlemezek eltávolítása után a preparátumokat 4×3'-ig 50% formamidot tartalmazó 1×SSC-vel, majd ugyancsak 4×3'-ig 2×SSC-vel mostuk 42°C-on. Szobahőmérsékletű 0.05% Triton X-100 tartalmú 4×SSC-vel (4×SSC/TX) történő öblítést követően a lemezekre 40μl fluoreszcein izo-tiocianáttal jelölt avidint (VECTOR, 500× hígítás 4×SSC/TX-ben) tettünk és lefedve 37°C-on nedveskamrában inkubáltuk őket 30-45'-ig. Ezután a lemezeket 3×2'-ig 4×SSC/TX-el mostuk 42°C-on. A mosást követően 40μl biotinnal jelölt anti-avidin ellenanyaggal (VECTOR, 100× hígítás 4×SSC/TX-ben) inkubáltuk a preparátumokat 30-45'-en at 37°C-on. Ezt a előzővel azonos 4×SSC/TX mosás, újabb réteg FITC-avidin, majd ismételt mosás követte.

**DNS festés, lefedés:** A kromoszómákat 1mg/ml parafenilén-diamint és 3μg/ml propidium jodidot tartalmazó PBS oldatban szobahőmérsékleten 5-10'-ig festettük. A preparátumokat ezután PBS-sel öblítettük, majd "anti-fading" oldattal (90% glicerol, 1mg/ml parafenilén-diamin) lefedtük. A fedőlemezzel történő lezárás után a preparátumokat fluoreszcens mikroszkóppal vizsgáltuk.



### 3.4. DNS technikák:

Az enzimeket a gyártó (*New England Biolabs, Amersham, Boehringer-Mannheim, Promega*) által ajánlott összetételű pufferekben és feltételek mellett használtuk. Az összes DNS manipulációt az ismert módszerek felhasználásával hajtottuk végre (88). A *Southern* hibridizációhoz próbaként használt DNS fragmenteket Megaprime kit (*Amersham*) használatával jelöltük. A CHO-K1 genomikus  $\lambda$  fág bank (*Stratagene*) screenelését a gyártó által ajánlott módon végeztük el. A di-deoxi szekvenálást egyszálú M13mp18-M13mp19 fág DNS-en manuálisan végeztük, vagy plazmid templáton automata szekvenátorral (*373 DNA Sequencing System, Applied Biosystem*) történt.

### 3.5. Szekvencia analízis:

A nukleotid szekvencia analíziseket a MicroGenie (*Beckman*) és PC/Gene programokkal végeztük. A homológia összehasonlítások BLAST-tal történtek (89).

### 3.6. Pulsed field gélelektroforézis:

**Agarózba ágyazott nagy molekulású emlős DNS preparálása:**  $2 \times 10^8$  CHO-K20 sejtet a tenyésztő csészékből összegyűjtöttünk. A sejteket PBS-ben felfuszpendáltuk majd  $10^4$ -ig 2000rpm-el centrifugáltuk. Ezután a pelletet 2.5ml PBS-ben szuszpendáltuk fel és a szuszpenzióhoz azonos térfogatú 1% LMP (*Low Melting Point agaróz PBS-ben felolvasztva*) agaróz oldatot adtunk. A keveréket megfelelő öntőformába töltöttük és azt az agaróz megszilárdulásáig  $4^{\circ}\text{C}$ -on tartottuk. Megszilárdulás után az agaróz darabkákat ESP pufferben (*0.5M EDTA pH:8, 1% Na-lauroil szarkozin, 100 $\mu\text{g}/\text{ml}$  Proteináz-K*) 24 órán at  $50^{\circ}\text{C}$ -on inkubáltuk. Az inkubációs idő letelte után az ESP puffert 50mM-os EDTA (*pH:8*) oldatra cseréltük le és a mintákat  $4^{\circ}\text{C}$ -on tároltuk.

**Az agarózba ágyazott DNS emésztése:** A nagy molekulású emlős DNS-t tartalmazó agaróz gélből  $\sim 5\mu\text{g}$  DNS-t tartalmazó darabkákat vágunk le. A Proteináz-K

inaktiválására az agaróz darabkákat 1mM PMSF-et tartalmazó TE pufferben 2 óráig inkubáltuk 37<sup>0</sup>C-on. A további ekvilibrálási lépések szintén 2 órán át tartottak 37<sup>0</sup>C-on. A PMSF kezelést követően TE pufferrel majd megfelelő restriktív enzim pufferrel ekvilibráltuk az agaróz darabkákat. A restriktív emésztés friss restriktív pufferben (~10U enzim/μg DNS) 12-18 órán át tartott a használt enzim számára ideális hőmérsékleten. Az emésztés befejezte után az agaróz darabkákat az elektroforézishez használt pufferrel öblítettük.

**Pulsed field gélelektroforézis paraméterei:** Az elektroforézist forgó tálcás *pulsed field* gélelektroforézis készüléken 1%-os agaróz gélen 0.25×TBE pufferben hajtottuk végre.

Az alkalmazott elválasztási paraméterek :

- Hőmérséklet: 15<sup>0</sup>C
- Elektromos térerősség: 5V/cm
- Ciklus idő: 20"
- Futási idő: 30 óra
- Forgási szög: 110<sup>0</sup>

### 3.7. PCR, RT-PCR:

**PCR:** A PCR reakciókat az Expand PCR kittel (*Boehringer-Mannheim*) hajtottuk végre a gyártó utmutatásai szerint. Az alkalmazott amplifikálási paraméterek a következők voltak:

- 94<sup>0</sup>C, 2'
- 10 ciklus: 94<sup>0</sup>C, 5"  
62<sup>0</sup>C, 30"  
67<sup>0</sup>C, 4' vagy 67<sup>0</sup>C, 2'
- 20 ciklus: 94<sup>0</sup>C, 5"  
62<sup>0</sup>C, 30"

67<sup>0</sup>C, 4'+20" minden ciklus után vagy 67<sup>0</sup>C, 2'+20" minden ciklus után

**RT-PCR:** Totál RNS-t CHO-K20 sejtekből a savas fenol módszerrel tisztítottunk (88). Az egy szálú cDNS szintézisét AMV reverz transzkriptáz kittel (*Boehringer-Mannheim*) végeztük a HPRF oligonukleotidot használva primerként. A PCR reakcióhoz a cDNS szintézis reakció felét (10ml) használtuk templátként. Primerként a HPRF és HPRR oligonukleotidokat használtuk.

**Használt oligonukleotidok:**

RUAF: GTGGCCTGAGAACCCCTCAGTAGACTCCACAGAC

RUAR: TGCCAAGCAACGAAGGAAGACTTGCCAGGAGCTA

HPRF: ATGGAAGGTCGCAGCCGTGTGGGAA

HPRR: TCAGGGGATTCTGTCCATGGGCTC

## **4. EREDMÉNYEK**

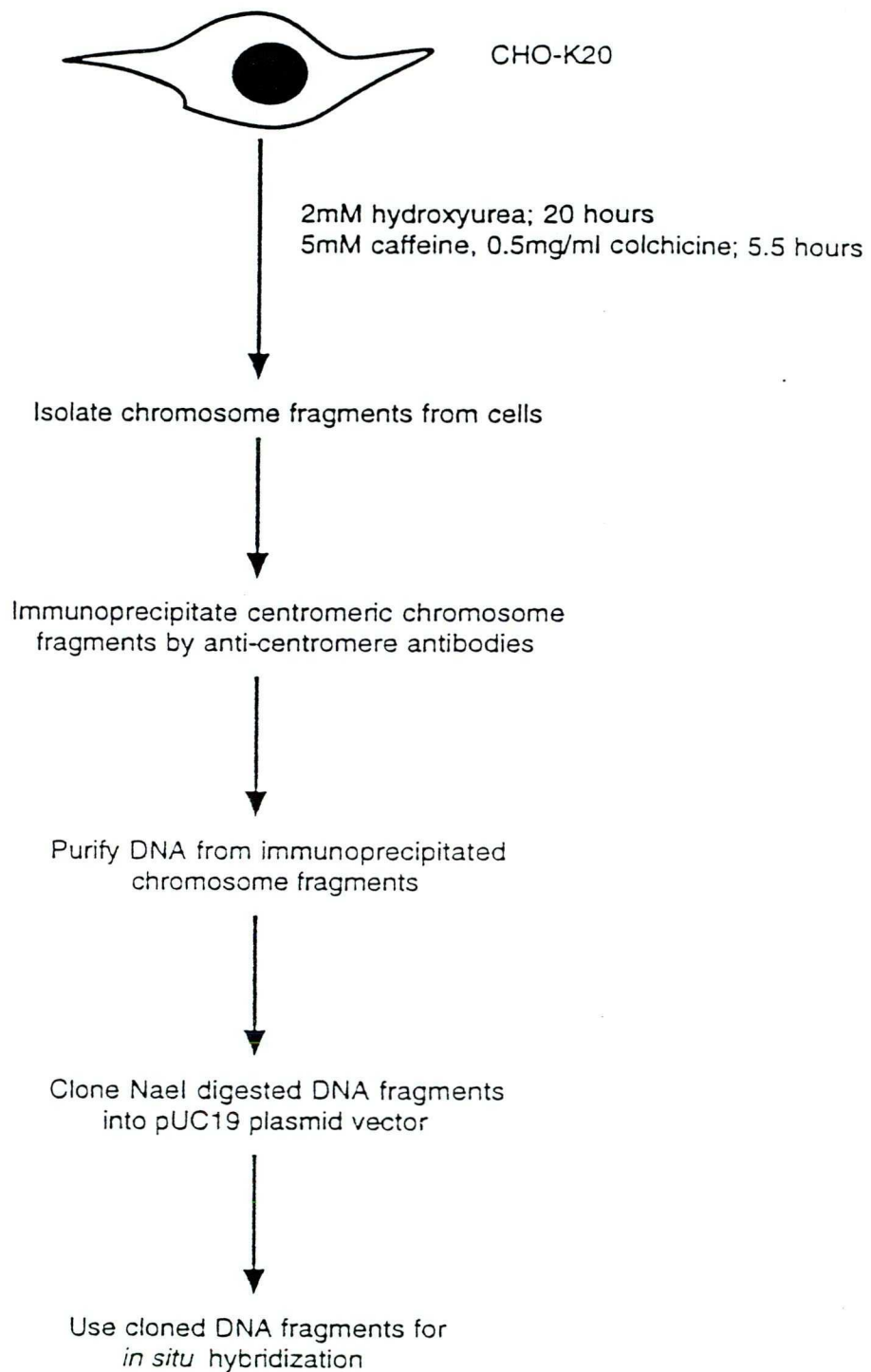
### **4.1. Kromoszóma specifikus centromer szekvencia izolálása kínai hörcsög sejtéből:**

Munkánk kezdeti célkitűzéseként megpróbáltunk eljárást kidolgozni centromerikus DNS elemekben dúsított kromatin frakció izolálására emlős sejtekből. Kísérleteink során arra a korábbi megfigyelésre támaszkodtunk amely szerint a hidroxüreával DNS szintézisükben gátolt hörcsög sejtek koffein kezelése a sejtek mitózis osztdását indukálja (90). A DNS szintézis elmaradása miatt a sejtekben normális morfológiájú kromoszómák helyett apró kromatin fragmentek láthatók. A centromerikus kromoszóma fragmentek centromer fehérjéket felismerő ellenanyagok használatával, immunofluorescenciás módszer révén azonosíthatók. Ily módon megfigyelhető, hogy ezek a kromoszóma darabok az intakt kromoszómákhoz hasonló módon a sejt ekvatoriális síkjába rendeződnek, majd normális anafázisos szegregációt mutatnak (91). A kinetochort hordozó centromerikus kromoszóma

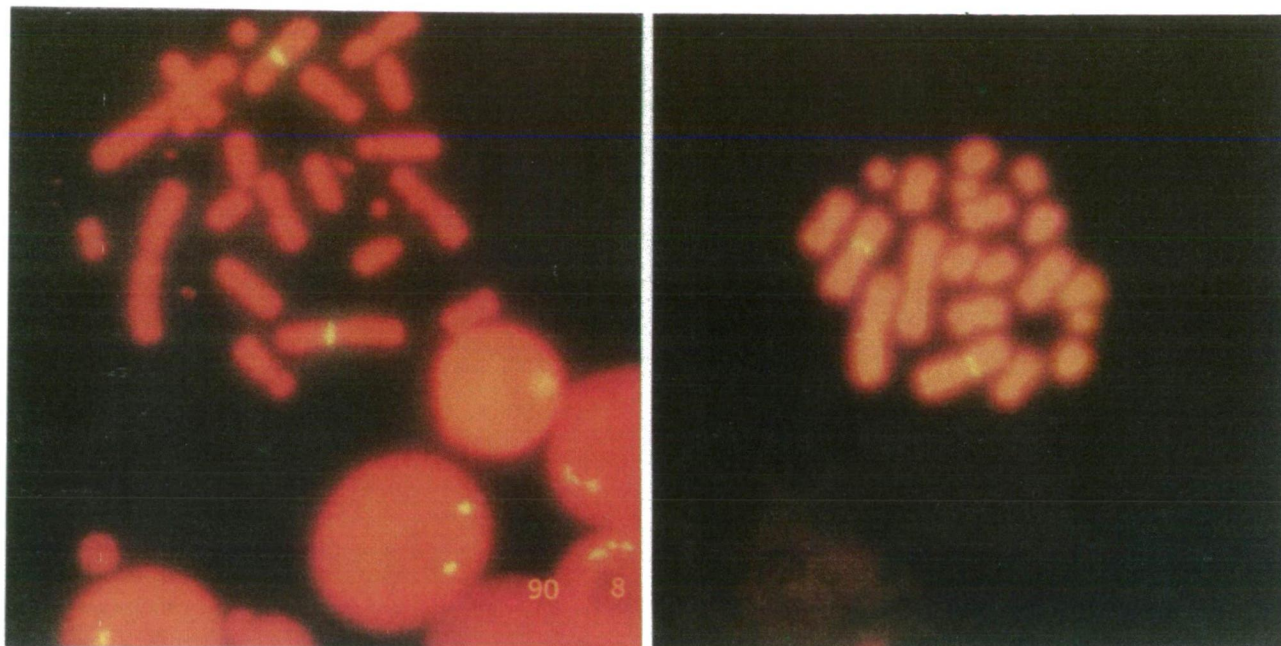


darabok látszólag normális mitotikus mozgásai arra utalnak, hogy a hidroxürea-koffein kezelés nem gátolja a kinetochornak és a magorsó mikrotubulusainak a megfelelő kapcsolódását. Ezzel összhangban vannak az elektronmikroszkópos vizsgálatok eredményei, amelyek az elvált centromer-kinetochor komplexek ép, az intakt kromoszómákra jellemző morfológiáját mutatták ki (80). Mindezek a megfigyelések felvették annak a lehetőségét, hogy a hidroxürea-koffein kezelést felhasználjuk a funkcionális centromer-kinetochor komplexet hordozó kromoszóma fragmenteknek a kromatin többségétől történő elválasztására, majd ezt követően megkíséreljük ezeknek a kromoszóma fragmenteknek a szelektív dúsítását centromer fehérjéket felismerő ellenanyagokkal végrehajtott immunoprecipitáció révén.

Az alkalmazott kísérleti eljárás vázlatát a 2. ábra mutatja.  $3-4 \times 10^8$  CHO-K20 (*Chinese hamster ovary*) sejtet 20 órán át hidroxürea jelenlétében növesztettünk. 20 óra után a táptalajt koffeinnel és a sejteket mitózisban blokkoló kolhicinnel egészítettük ki. A néhány óra múltán begyűjtött mitotikus sejtekből a kromoszóma darabokat detergens kezeléssel kiszabadítottuk, majd a centromerikus kromoszóma fragmenteket immunoprecipitáltuk centromer fehérjéket felismerő ellenanyagok felhasználásával. Az immunoprecipitált kromoszóma darabokból tisztított DNS fragmenteket plazmid vektorba klónoztuk. A kapott plazmidok közül nyolc, hosszabb inzertet tartalmazó klónt kiválasztottunk és a klónozott fragmentek lokalizációját metafázisos kínai hörcsög kromoszómákon *in situ* hibridizációval vizsgáltuk. Egy klón, HUCAFF170 (*HUCAFF: hydroxyurea-caffeine*) erős hibridizációt mutatott két kromoszóma pericentrikus régiójához CHO-K20 sejtben (3a. ábra). Az élő állatból származó lépsejteken megismételt *in situ* hibridizáció révén a hibridizáló kromoszóma párt a kínai hörcsög 2. kromoszómájaként azonosítottuk (3b. ábra).



**2. ábra. Kínai hörcsög kromoszómák centromerikus kromatin frakciójának izolálására alkalmazott eljárás vázlata.**



a

b

**3. ábra. A HUCAFF170 próba hibridizációja CHO-K20 sejtek (a) és kínai hörcsög lépsejtek (b) metafázisos kromoszómáihoz.**

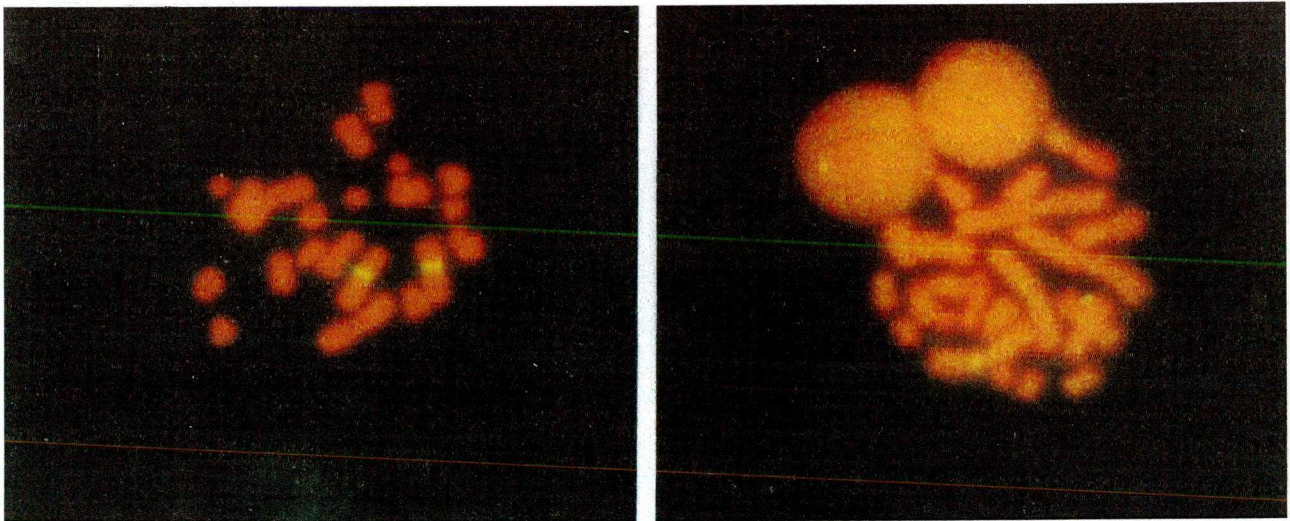
**4.2. Az azonosított centromerikus DNS fragment egy új, faj és kromoszóma specifikus repetitív DNS elem része:**

A HUCAFF170 klón szekvenca analízise kimutatta, hogy az 536 bp hosszú fragment nem mutat homológiát ismert DNS elemekkel. A DNS fragment fajspecifitását *Southern* és *in situ* hibridizációval vizsgáltuk meg. A humán és egér kromoszómákon végrehajtott *in situ* hibridizáció során specifikus hibridizációs szignált nem detektáltunk. A HUCAFF170 klónt próbaként használva humán, egér és szíriai hörcsög genomikus DNS-en végrehajtott *Southern* hibridizáció ugyancsak negatív eredményt hozott. Mindezek a megfigyelések arra utaltak, hogy a HUCAFF170 klón egy evolúciósan nem konzervált, valószínűleg csak a kínai hörcsögben előforduló fajspecifikus DNS elem.



Az *in situ* hibridizáció során megfigyelt szignál intenzitása arra utalt, hogy a HUCAFF170 fragment a 2. kromoszóma pericentrikus régiójában több példányban fordul elő. Az 536 bp hosszú fragment szekvenciája belső periodicitást nem mutatott. Ez valószínűvé tette, hogy a klónozott fragment egy hosszabb repetitív DNS elem része lehet. Ezt a feltételezést a HUCAFF170 klónnal homológiát mutató hosszabb DNS szakasz szerveződésének vizsgálatával kívántuk ellenőrizni.

Az eredetileg azonosított 536 bp-os fragmentet próbaként használva kínai hörcsög DNS bankból izoláltunk egy 13.3 kb méretű genomikus inzertet tartalmazó lambda fág klónt (HC90). *In situ* hibridizáció során a HC90 klón a HUCAFF170 fragmettel azonos módon a 2. hörcsög kromoszóma pericentrikus régiójához hibridizált (4a. ábra).



a

b

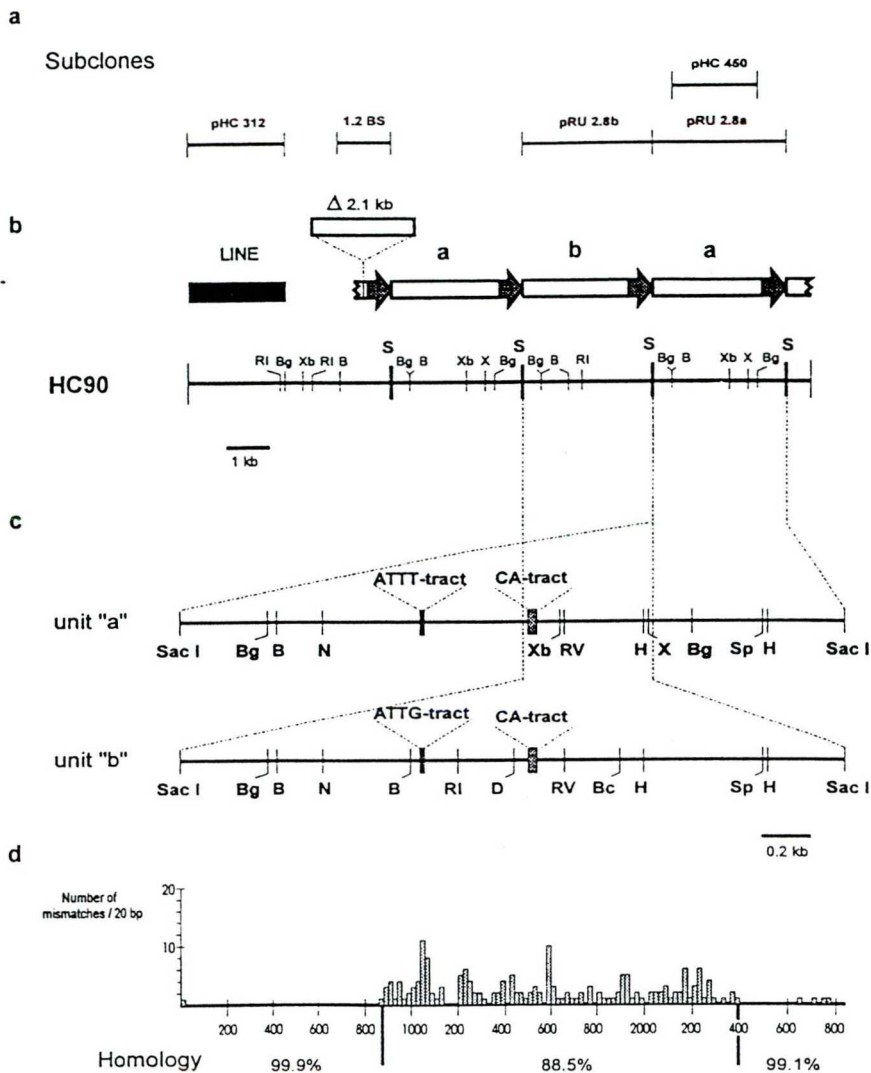
**4. ábra. A HC90 (a) és a pHC312 (b) próbák hibridizációja kínai hörcsög lépsejtek metafázisos kromoszómáihoz.**

A restrikciós térképezés során kimutattuk, hogy a fág klón három, 2.8 kb méretű tandem elrendeződésű SacI fragmentet tartalmaz, amelyek hibridizálnak a HUCAFF170 fragmenttel (5. ábra). A három SacI fragment közül a két szélső azonos restrikciós mintázatot

mutatott. Ezen SacI fragmentek középső, 1.8 kb méretű BglIII szubfragmentjét plazmid vektorba klónoztuk (pHC450, 5a. ábra). Habár a pHC450 klón nem hibridizált a HUCAFF170-el, *in situ* hibridizációban próbaként használva ez a klón is kizárólag a 2. kromoszóma pericentrikus régiójában adott jelet. A középső SacI fragment restrikciós mintázata jelentősen eltért a két szélső egységtől. Ennek ellenére ez a szakasz is hibridizált mind a HUCAFF170 és a pHC450 klónokkal. Összefoglalva ezek az adatok arra utaltak, hogy az eredetileg azonosított 536 bp méretű HUCAFF170 fragment részét képezi egy 2.8 kb méretű ismétlődő egységgel jellemezhető, kromoszóma specifikus, tandem repetitív DNS elemnek. A HC90 fág három ilyen ismétlődő egységet tartalmaz. A két azonos restrikciós mintázatot mutató "a" típusú egységet egy eltérő restrikciós mintázatú repetitív egység variáns választja el ("b" típusú repetitív egység).

A fág klón a 3' végén egy 0.5 kb méretű fragmentben végződik, amely hibridizál az ismétlődő egységgel és arra utal, hogy a fágban klónozott genomikus DNS szakasz a kromoszómán ebben az irányban az azonosítottakhoz hasonló repetitív egységekben folytatódik. A HC90 klón 5' végi 4.4 kb-os szakaszának része, egy közvetlenül az "a" típusú egység szomszédságában elhelyezkedő 1.2 kb méretű BamHI-SacI fragment, amely ugyancsak tartalmaz a HUCAFF170 klónnal homológ szekvenciákat. A szekvencia analízis kimutatta, hogy ez a szakasz egy deléciós ismétlődő egységnek felel meg (lásd később). A legszélső 3.2 kb-os SalI-BamHI fragment nem hibridizál a 2.8 kb-os SacI repetitív egységgel. Ezen régió 2.1 kb méretű SalI-BglIII szubfragmentjét plazmid vektorba klónoztuk (pHC312), majd próbaként használtuk *in situ* hibridizációhoz. A pHC312 klón egyenletesen hibridizált az összes kínai hörcsög kromoszómához a karok teljes hosszában (4b ábra). Ez a hibridizációs kép arra utalt, hogy a próbaként használt fragment szétszórt repetitív egységekkel homológiát mutató szakaszokat tartalmaz. A pHC312 klón későbbi szekvencia analízise valóban kimutatta, hogy ez a DNS szakasz egy az emlősökben általánosan előforduló retroponon eredetű LINE elemnek felel meg.





5. ábra. (a) A HC90 fág néhány szubklónjának a szerkezete. (b) A HC90 fág restrikciós térképe. A szürkével jelzett szakaszok homológok a HUCAFF170 klónnal. A fekete szakasz a LINE elem elhelyezkedését mutatja. (c) Az "a" és "b" típusú repetitív egységek restrikciós térképei. (d) Az "a" és "b" típusú repetitív egységek szekvencia különbségeinek eloszlása.

#### 4.3. Az "a" és "b" típusú ismétlődő egységek szekvencia analízise:

Az "a" és "b" típusú repetitív egységeknek megfelelő 2.8 kb méretű SacI fragmenteket M13 fág vektorba szubklónoztuk, majd elvégeztük azok szekvencia analízisét (1., 2. melléklet). A 2847 bp hosszú "a" elem 93.6% hasonlóságot mutatott a 2843 bp hosszú



"b" típusú elemmel. Az HUCAFF170 fragment 5' végi 369 bp-nyi szakasza 90.6%-ban homológ a Sac I elemek 3' végével. A maradék 167 bp méretű szakasz nem mutatott homológiát ismert DNS szekvenciákkal. A SacI elemek bázis összetétele összességében kiegyensúlyozott (G-C: 52.1%), habár a 3' és 5' végek enyhén G-C gazdagok, míg a középső néhány száz bázispárnyi szakasz (különösen 1100-1500 között) magas A-T tartalmú. Az egész elemre kiterjedő rövidebb belső periodicitást nem találtunk. A szekvencia összehasonlítások során az ismétlődő egységek nem mutattak kiterjedtebb homológiát más ismert DNS szekvenciákkal, kivéve két rövid régiót (2470-2620 és 1-73) amelyek ~70 illetve ~65% homológiát mutattak homeotikus gének DNS kötő homeobox szakaszával (lásd még később).

#### ***4.3.1. Az ismétlődő egységek különleges DNS elemeket tartalmaznak:***

A repetitív egységek középső szakaszában több különleges DNS elem koncentrációja figyelhető meg. A magas A-T tartalom mellett rövid szekvencia részletek tandem ismétlődéséből felépülő ún. "egyszerű" szakaszok azonosíthatók. Ilyen szekvencia elem például az "a" típusú egységben azonosítható (ATTT)<sub>10</sub> szakasz (1051-1091). A tíz tetranukleotid egység közül kilenc tökéletesen ismétlődik, míg egy esetben T→C tranzíció látható. A "b" típusú elemben ugyanezt a pozíciót egy (ATTG)<sub>6</sub> "egyszerű" szakasz foglalja el, ahol is a hat ismétlődő egység közül kettő tartalmaz T→A transzverziót. Az "a" és "b" típusú elemekben ekvivalens pozíciót elfoglaló, különböző "egyszerű" szekvenciák evolúciós kapcsolatának vizsgálata érdekes következtetésekhez vezethet. Valószínűtlennek tűnik, hogy a tetranukleotid szakaszok amplifikációját követően minden egyes tetranukleotid elem utolsó nukleotidja egymástól függetlenül T-ről G-re (vagy G-ről T-re) cserélődött az egyes SacI elemekben. Az utolsó nukleotid cseréjének a tetranukleotid egység amplifikációja előtt kellett megtörténnie. Ez maga után vonja, hogy a tetranukleotid egységekből felépülő "egyszerű" szakaszok a 2.8 kb-os repetitív elemekben lokálisan jöttek létre. Ez két alternatív módon

játszódhatott le: (i) Feltételezik, hogy az "egyszerű" szekvenciák a DNS replikációja során szál elcsúszások (*polimerase slippage*) révén alakulnak ki. Ha a SacI elemekben az "egyszerű" szakaszok kialakulása ugyanazon pozícióban egymástól függetlenül történt, feltételezhetően a tetranukleotid helyzete az amplifikáció szempontjából kitüntetett volt. Lehetséges, hogy a tetranukleotidot határoló szekvenciák hajlamosítottak az amplifikációhoz vezető szál elcsúszásokra, vagy egyéb nem tisztázott módon indukálhatták a tetranukleotid szakaszok helyspecifikus amplifikációját. (ii) Ugyancsak lehetséges magyarázat, hogy a különböző összetételű "egyszerű" szakaszok a genom más részein alakultak ki és a SacI elemekbe helyspecifikus integráció révén épültek be. A helyspecifikus integráció maradványai lehetnek az "a" és "b" elemek "egyszerű" szakaszait határoló pirimidin gazdag, rövid direkt ismétlődések.

Mindkét repetitív egység tartalmaz egy hosszú, alternáló C-A nukleotidokból felépülő szakaszt ("a" elem: 1524-1582, "b" elem: 1509-1583). Ezt az "egyszerű" szekvencia elemet megtalálták az eddig vizsgált összes eukariótában. Az ilyen, Z-DNS struktúra felvételére képes DNS szakaszokról feltételezik, hogy rekombinációs "hot spot"-ként működhetnek és ezáltal fontos szerepet tölthetnek be tandem repetitív szekvenciák evolúciójában (92).

A SacI egységek 1100-1500 bp közötti régiója további rövid, tandem ismétlődő szekvencia elemeket és  $d(A-T)_{n>3}$  szakaszokat tartalmaz. Korábbi kísérleti adatok azt mutatják, hogy hosszabb, kizárólagosan A-T nukleotidokból álló szakaszok jelentős hatással vannak a kromatin szerkezetére. Az ilyen szekvencia elemek a DNS görbülését válthatják ki, ez viszont az adott DNS szakaszon elhelyezkedő nukleoszómák pozícióját befolyásolhatja (93). Más megfigyelések az A-T gazdag szakaszoknak az amplifikációban betöltött potenciális szerepét huzzák alá (94).



#### 4.3.2. A "b" típusú repetitív elem valószínűleg gén-konverzió eredménye:

A két vizsgált SacI elem szekvenciájának összehasonlítása a mutációk asszimmetrikus eloszlását mutatta a repetitív egységek mentén (5d. ábra). Habár a két elem szekvenciája 93.6%-ban azonos, a középső ~1.5 kb-os szakasz ennél nagyobb fokú eltérést mutat (88.5% homológia), míg az 5' és 3' végi régiók gyakorlatilag azonosak (99.9% illetve 99.1% homológia). A mutációk ilyen fokú asszimmetrikus eloszlásának legvalószínűbb magyarázata az, hogy a "b" elemet, egy "a" típusú egység és egy szekvenciájában jelentősen eltérő repetitív egység variáns között lejátszódott kettős *crossing over* majd ezt követő gén konverzió hozta létre. Nem tudjuk azonban azt sem kizárni, hogy az 5' és 3' végi szakaszok látszólagos konzervativitása valamilyen funkcióval kapcsolatos szelekciós nyomás eredménye.

#### 4.3.3. Az "b" és "a" elemek határán egy hosszú, kettős homeobox szakaszt tartalmazó nyílt leolvasási keret található:

A 2.8 kb-os SacI elemek szekvenciáinak analízise nem mutatott ki fehérjék kódolására alkalmas, hosszabb nyílt leolvasási kereteket (*open reading frame: ORF*). Ha azonban az "a" és "b" típusú elemeket b-a sorrendben egymás mögé illesztjük az egységek határán egy 1062 bp hosszú nyílt leolvasási keret jelenik meg (6a. ábra). A potenciálisan 353 aminosavból álló polipeptidet kódolni képes ORF két hosszabb szakaszt tartalmaz amelyek DNS szinten ~65% illetve ~70% homológiát mutatnak a homeotikus gének erősen konzervált homeobox régiójával. A homológia fehérje szinten is megfigyelhető. A hosszabb, 60 aminosavból álló szakasz (17-76) esetében az aminosavak 45%-a azonos a patkány LIM-2 homeobox fehérjéjének DNS kötő homeodoménjével. Ha figyelembe vesszük a konzervatív aminosav cseréket is, a hasonlóság 65%. A második, rövidebb szakaszon (139-165) az alkotó aminosavak 55%-a azonos a LIM-2 homeodoménjével, de a konzervatív aminosav cserékkel

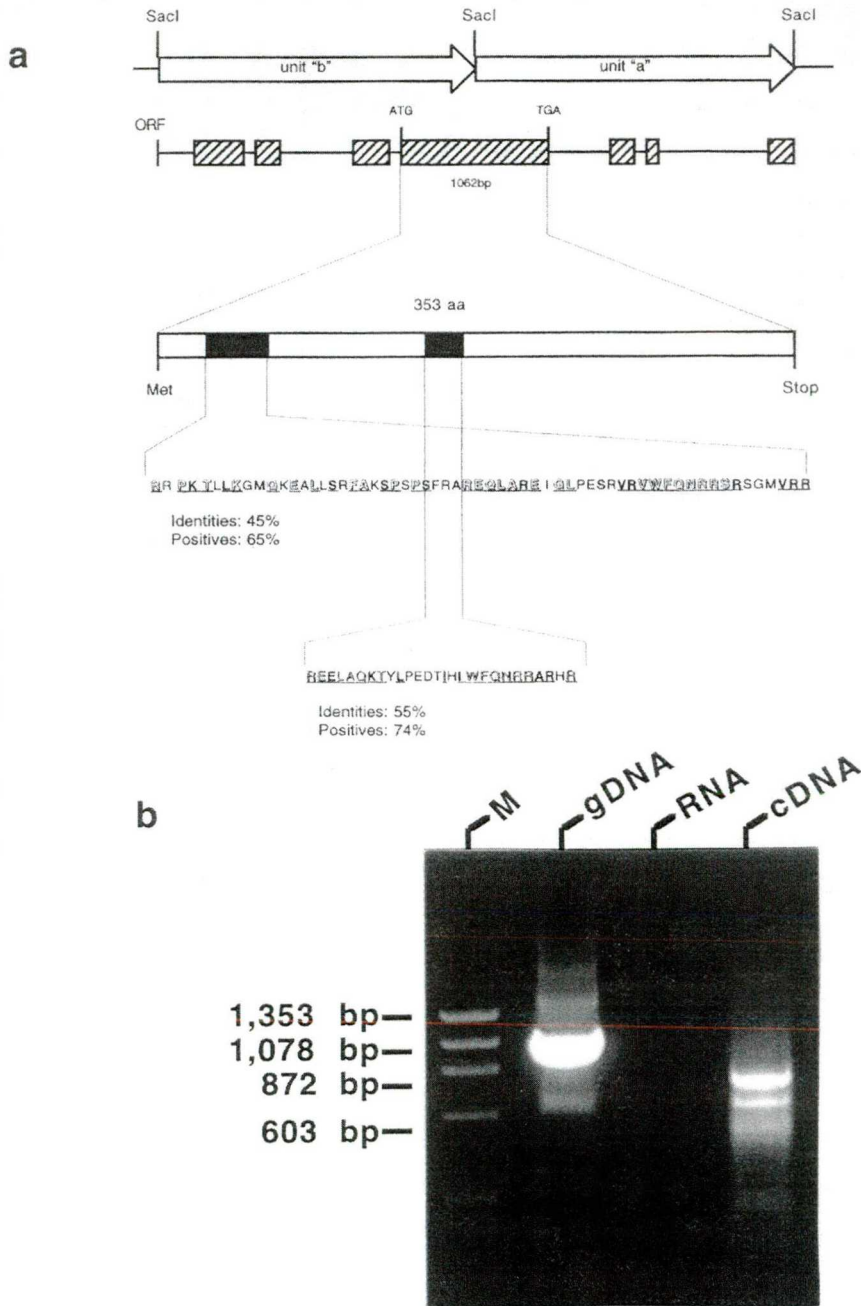


együtt a hasonlóság ez esetben is jóval nagyobb (74%). Hasonló fokú homológia más gerincesekből sőt rovarokból származó homeotikus génekkel is megfigyelhető.

Az azonosított nyílt leolvasási keret expresszióját RT-PCR-rel vizsgáltuk meg. CHO-K20 sejtekből tisztított totál RNS-ről a leolvasási keret elejére tervezett oligonukleotid primert alkalmazva először egyszálú cDNS-t szintetizáltunk. Ezt követően a cDNS-t templátként használva, a leolvasási keret elejére és végére specifikus oligonukleotid primer párral PCR reakciót hajtottunk végre. Ily módon két specifikus DNS fragmentet tudunk amplifikálni amelyek azonban a vártnál kisebbnek bizonyultak (~700 bp illetve ~900 bp, 6b. ábra). A továbbiakban a PCR fragmenteket klónoztuk és elvégeztük azok részleges szekvencia analízisét. A vizsgált 5 klón közül 3 azonosnak bizonyult. Habár az összes esetben a primerekkel homológ szakaszokat a klónok elején és végén azonosítani tudtuk, a továbbiakban a fragmentek nem mutattak homológiát sem a repetitív elemmel sem más ismert DNS szekvenciával. A nyílt leolvasási keret expresszióját tehát jelenleg nem tudjuk egyértelműen eldönteni. Lehetséges, hogy az RT-PCR-rel amplifikált fragmentek a repetitív elembe ágyazott expresszálandó DNS szakaszokat reprezentálnak. Azt sem tudjuk azonban kizárni, hogy az alkalmazott primerekkel homológ szakaszok előfordulhatnak más, nem rokon expresszálandó DNS elemekben is.

Az itt leírt kettős homeoboxot tartalmazó repetitív szekvenciához hasonló DNS elemet emberben is leírtak (24). A 4. humán kromoszómán lokalizálódó, 3.3 kb méretű ismétlődő egységgel jellemezhető, tandem szerveződésű D4Z4 repetitív elemet egy autoszómás domináns öröklődő betegség (*fascioscapulohumeral muscular dystrophy*: FSHD) kapcsán azonosították. A betegség kialakulásának pontos mechanizmusa még nem ismert, de a defektust okozó mutáció minden esetben a repetitív elemek deléciójával jár. A D4Z4 elemmel homológ szekvenciák megtalálhatók emberszabású és egyéb óvilági majmokban is (94). Habár az általunk azonosított repetitív szekvencia a kettős homeoboxtól eltekintve nem

mutat szekvencia homológiát a D4Z4 elemmel a szerveződésbeli hasonlóságok mégis figyelemre méltóak.



6. ábra. A HC2 szatellit repetitív egységeiben található nyílt leolvasási keretek elhelyezkedése (a). A homeodoménnal homológiát mutató szakaszok feketével jelöltek. A patkány LIM-2 homeobox fehérjéjével azonos aminosavakat üres, míg a konzervatív cseréket fekete aláhúzott betűk jelölik. (b) A leghosszabb ORF expressziójának vizsgálata RT-PCR-rel.

#### **4.4. Az 1.2 kb méretű BamHI-SacI fragment egy deléciós repetitív egységet tartalmaz:**

Az "a" típusú repetitív egység szomszédságában elhelyezkedő, a HUCAFF170 klónnal hibridizáló 1.2 kb méretű BamHI-SacI fragment szekvencia analízisét elvégeztük (3. melléklet). Az 1164 bp-os fragment két különböző régióból tevődik össze. Az "a" elemmel közvetlenül határos 617 bp hosszú szakasz 99.2% illetve 98.2% homológiát mutat az "a" illetve "b" típusú elemek 3' végével. Az "a" elemtől távolabb eső A-T gazdag szakasz nem mutat homológiát a repetitív egységekkel. Ezt a két régiót egy olyan 66 bp méretű fragment választja el, amely 86.4%-ban hasonló a SacI elemek 130-195 bp közötti szakaszához. Mindezeket úgy értelmezhetjük, hogy az 1.2 kb-os BamHI-SacI fragment egy olyan csonka repetitív egységnek felel meg amelynek hiányzik az első 129 bp-nyi valamint a középső 2.1 kb-nyi szakasza.

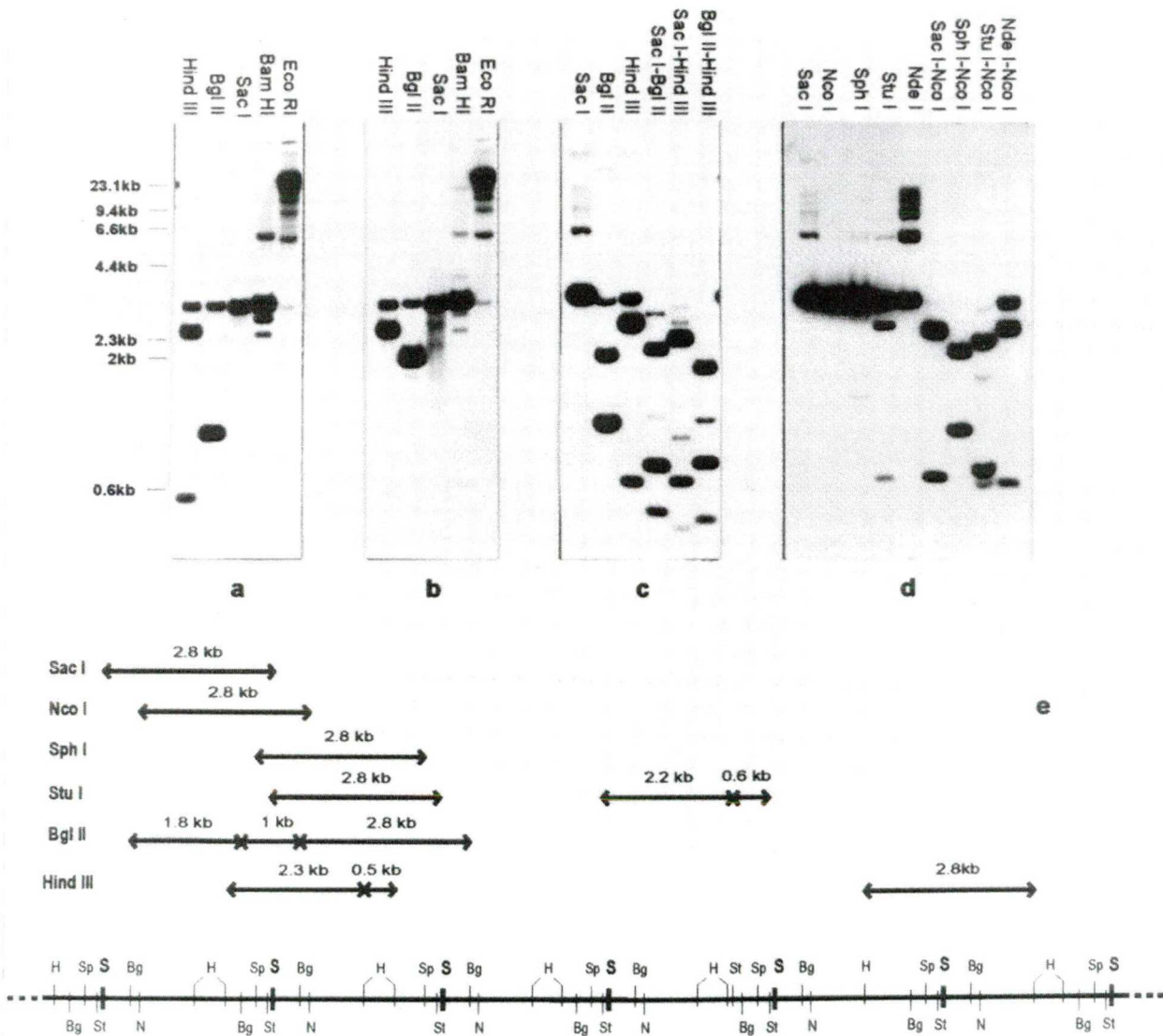
#### **4.5. Az új repetitív elem a 2. kínai hörcsög kromoszóma pericentrikus régiójának az egyik fő szatelliteszerű alkotó eleme:**

A következőkben megvizsgáltuk, hogy az új repetitív DNS elem HC90 klónban megfigyelt tandem szerveződése általánosítható-e az elemmel homológ genomikus szekvenciák többségére. CHO-K20 sejtekből tisztított genomikus DNS-t restriktív enzimekkel emésztettük, a fragmenteket gélelektroforézissel agaróz gélen elválasztottuk, membránhoz rögzítettük, majd a repetitív elemre specifikus radioaktívan jelölt DNS fragmentekkel hibridizáltuk. A kísérletek eredményét a 7. ábra mutatja.

A HUCAFF170 és pH450 fragmenteket próbaként használva hasonló restriktív mintázatokat kaptunk (7a. és 7b. ábra) ami arra utalt, hogy az alkalmazott próbákkal homológ szakaszok többsége azonos restriktív fragmenteken helyezkednek el. Kettős restriktív emésztések megerősítették, hogy a repetitív elemmel homológ szekvenciák valóban a HC90 klónban azonosítottakhoz hasonló ismétlődő egységek részeként fordulnak elő a genomban (7c. ábra). A 2.8 kb-os elemek tandem, szatelliteszerű szerveződésére utalt az

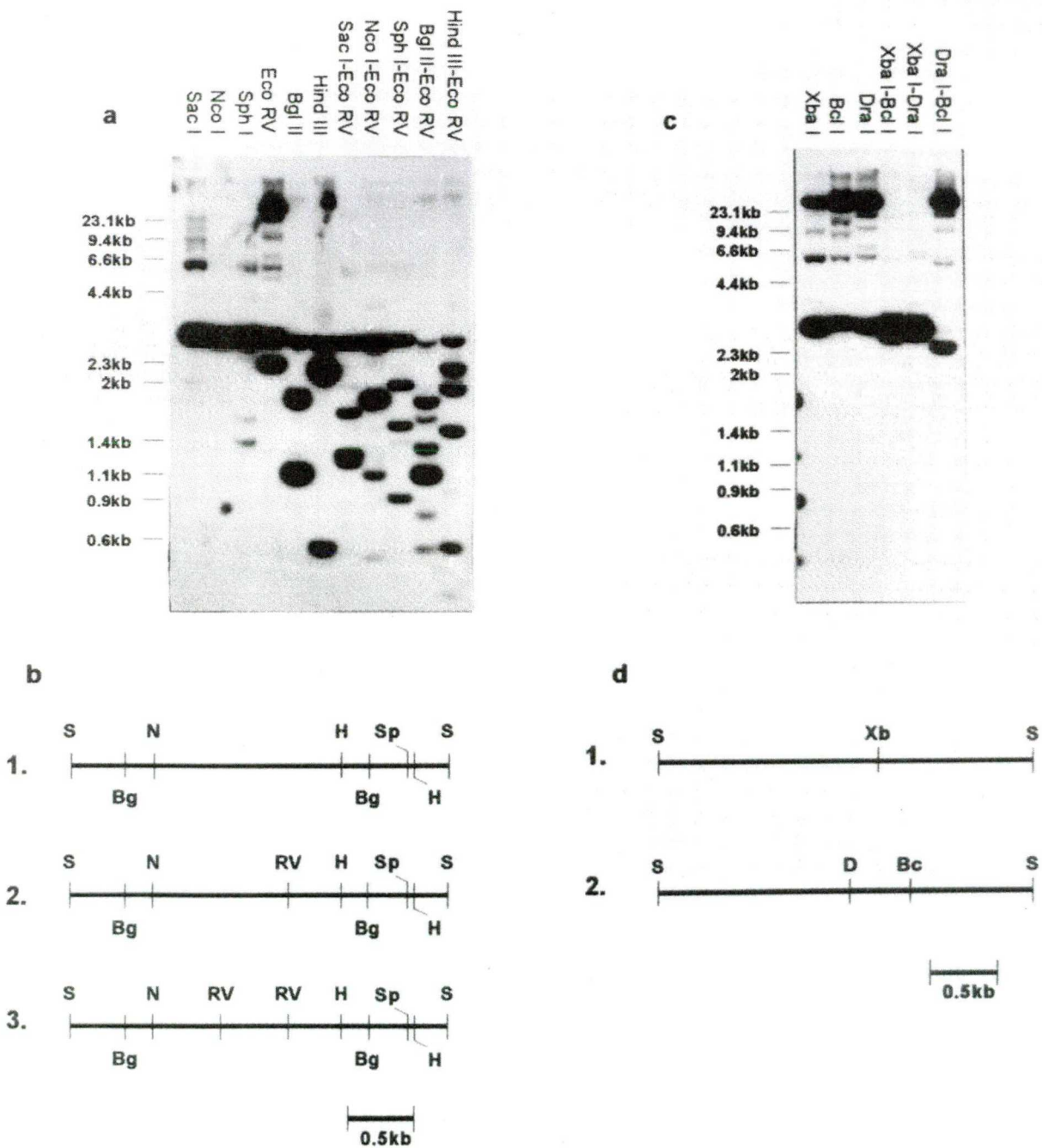


EcoRI, SacI és NdeI enzimekkel történő emésztés esetén látható létraszerű fragment mintázat (7d. ábra). Ugyancsak megfigyelhető, hogy a repetitív elemek többsége a HC90 klón "a" típusú egységeihez hasonló restrikciós mintázatot mutat (7e. ábra).



7. ábra. A HC2 szatellit szerveződésének vizsgálata CHO-K20 sejten Southern hibridizációval. Az alkalmazott próbák a következők voltak: (a) HUCAFF170, (b) pHC450, (c, d) pRU2.8b. A HC2 szatellit kisléptékű genomikus szerveződése (e).

A repetitív egységek középső részeinek jelentős heterogenitását detektáltuk az ebben a régióban vágó restriktív enzimekkel (8. ábra). Az EcoRV enzim felismerő helyére nézve három fő repetitív egység variáns léte kimutatható (8a. ábra). Az XbaI, DraI és BclI helyek eloszlása két fő, az "a" illetve "b" egységekhez hasonló mintázatot mutat (8b. ábra). Ezek a megfigyelések aláhúzzák a repetitív egységek középső szakaszainak és végeinek eltérő konzervativitását.



8. ábra. A HC2 szatellit repetitív egység variánsainak analízise Southern hibridizációval

(a, c). Az EcoRV (b) és XbaI (d) repetitív egység variánsok szerkezete.

Az repetitív elemek kópia számának meghatározásához ismert mennyiségű klónozott SacI elem és ismert mennyiségű CHO-K20 genomikus DNS hibridizációs intenzitásait hasonlítottuk össze, a repetitív elemre specifikus fragmentet használva próbaként. Így módon elvégezve a kópia szám becslést a repetitív egységek száma 2500-5000 közöttinek adódott haploid genomként. Figyelembe véve a 2.8 kb-os SacI elemek tandem szerveződését, az új repetitív elem 7-14 Mb-nyi DNS szakaszt foglalhat el, miáltal a 2. kromoszóma pericentrikus régiójának az egyik fő szatellit-szerű komponense lehet (HC2 szatellit).

#### **4.6. A HC2 szatellit nagy méretű homogén doméneket alkot:**

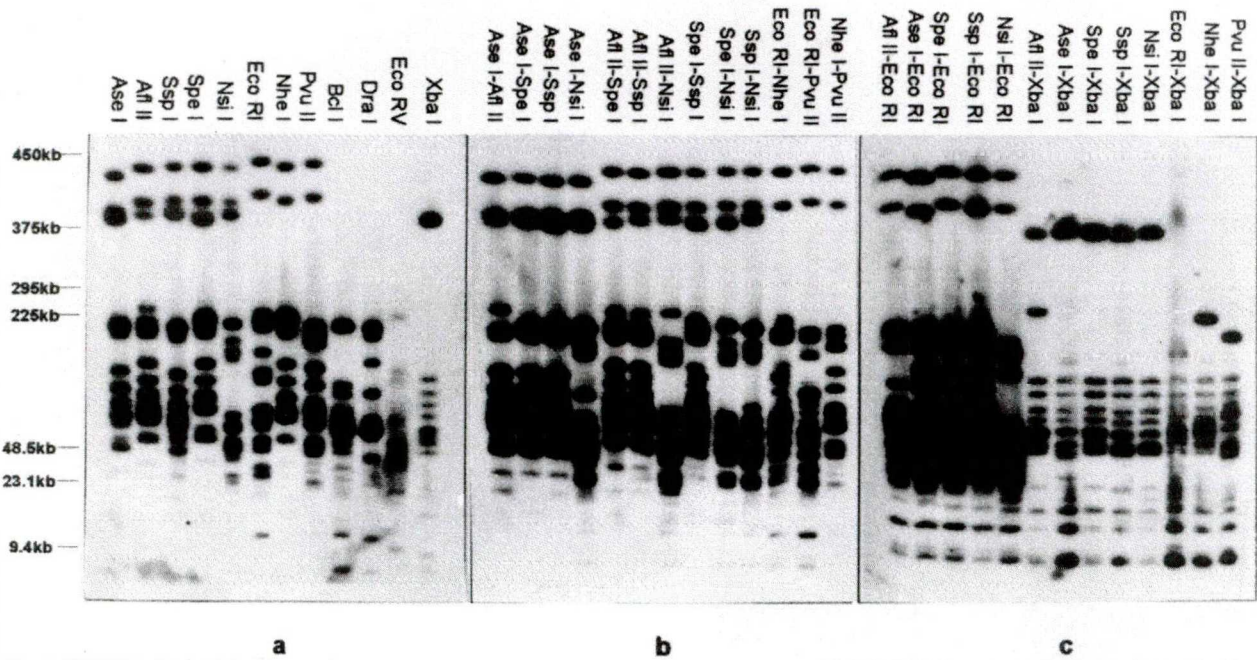
A HC2 szatellit nagy léptékű genomikus szerveződését változó erőterű gélelektroforézissel (*Pulsed Field Gel Electrophoresis*: PFGE) vizsgáltuk meg. A PFGE módszere lehetővé teszi olyan DNS fragmenteknek a méretük alapján történő szeparálását, amelyeket a konvencionális gélelektroforézis módszerével már nem lehet elválasztani.

CHO-K20 sejtekből nagy molekulásúlyú genomikus DNS-t preparáltunk. A restriktív enzimekkel emésztett DNS-t PFGE-vel elválasztottuk, majd a HC2 szatellittel homológ DNS szakaszokat radioaktívan jelölt pRU2.8b fragmenttel detektáltuk. A *Southern* hibridizációk eredményét a 9. ábra mutatja. Az alkalmazott restriktív enzimek nagy része meglepően hasonló fragment mintázatot adott. A 12 restriktív enzim közül 9 eredményezett fragmenteket a 370-450 kb mérettartományban, míg a homológ szekvenciák maradéka 260 kb-nál kisebb fragmenteken lokalizálódott. A 370-450 kb tartományban detektálható fragmentek száma alapján a használt enzimek négy csoportot alkotnak. Az AseI, AflII, SspI, SpeI és NsiI enzimek három fragmentet adnak az említett mérettartományban, amelyek a tíz lehetséges kombinációban elvégzett kettős emésztések után is megmaradnak. A kettős emésztések után megmaradó hasonló fragment mintázatok annak feltételezésével értelmezhetők, hogy a használt enzimek felismerő helyei egymáshoz közel, legfeljebb néhány kilobázispárnyi távolságra vannak. A második csoportot alkotó EcoRI, NheI és PvuII két,



370-450 kb közötti fragmentet eredményez. A csoporton belüli három lehetséges kettős emésztés a fragment mintázatot nem változtatja meg jelentősen, ugyanakkor az EcoRI és az első csoport enzimeinek kettős emésztései az első csoportra jellemző 375 kb-os fragment eltűnését eredményezik. A harmadik csoportot egymaga alkotó XbaI enzim egy 370 kb méretű intenzív sávot, valamint néhány, 90 kb-nál kisebb, halvány fragmentet szolgáltat. Az XbaI és az első csoport restriktív enzimeinek kettős emésztése során minden esetben megmarad a 370 kb körüli fragment ami arra utal, hogy ez a fragment ekvivalens az első csoport enzimeit által eredményezett ~375 kb méretű fragmenttel. Összhangban van ezzel a feltételezéssel az, hogy ugyanakkor a második csoport enzimeit (amelyek nem adnak 375 kb-nál fragmentet) a 370 kb-os XbaI sáv eltűnését eredményezik. Az AflIII-XbaI kettős emésztés további fontos adatokat szolgáltat. Az első csoportba tartozó AflIII a vártan megfelelően meghagyja a 370 kb-os XbaI sávot, viszont annak intenzitása jelentősen csökken és ugyanakkor egy erős fragment jelenik meg 280 kb körül. Ez a kép csak úgy értelmezhető ha feltételezzük, hogy az intenzív 370 kb-os XbaI sáv legalább kettős, méretében nagyon hasonló fragmentnek felel meg. Az alkalmazott enzimek negyedik csoportja (BclI, DraI, EcoRV) nem ad fragmentet a 370-450 kb-os tartományban.

Összefoglalva, a PFGE térképezés eredményei azt mutatják, hogy a HC2 szatellit több nagyméretű homogén blokkot alkot a genomban. A blokkok között több különböző restriktív enzim felismerő helyei koncentrálnak. Jelenleg még nem tudjuk, hogy ezek a blokkok közötti szakaszok a HC2 szatellit heterogén repetitív egységeiből, vagy esetleg más, nem rokon DNS elemekből épülnek-e fel. A homogén blokkok kialakulása a legkézenfekvőbb módon egy vagy néhány repetitív egységnek a lokális, szaltatorikus amplifikációjával magyarázható. Az azonos méretű, 370 kb-os XbaI fragmentek létrejöttét ugyanakkor a testvér kromatidák között lejátszódó egyenlőtlen *crossing over* eredményezhette.



9. ábra. A HC2 szatellit nagyléptékű szerveződésének vizsgálata pulsed field gélelektroforézissel.

#### 4.7. A HC2 szatellittel homológ DNS elemek fordított ismétlődéseket tartalmaznak:

A HC2 szatellit genomikus szerveződésének vizsgálata során egyértelműen bizonyítottuk, hogy az új repetitív DNS elemmel homológ szekvenciák túlnyomó többsége tandem ismétlődő 2.8 kb-os egységek formájában található meg a kínai hörcsög 2. kromoszómáján. A *Southern* hibridizációk során azonban gyakran láthatók kevésbé intenzív sávok is amelyek vagy a HC2 szatellit és egyéb DNS elemek határáról vagy egyéb módon átrendeződött régiókból származhatnak. Ezeknek az átrendeződött szakaszoknak a természetéről szolgáltatott adatokat az egy primerrel végrehajtott PCR kísérletek. Konvencionális PCR reakció során a DNS komplementer szálaira tervezett, két szekvencia specifikus oligonukleotid primer közötti DNS szakaszt amplifikáljuk. Egy primeres PCR reakcióban csak abban az esetben kaphatunk specifikus terméket ha a primerrel komplementer, egymás közelében elhelyezkedő, fordítottan ismétlődő szekvenciák találhatóak a templátként használt DNS-ben. Ilyen, a HC2 szatellittel rokon fordított ismétlődéseket próbáltunk keresni a 2.8 kb-os repetitív egység különböző részeire tervezett primerek

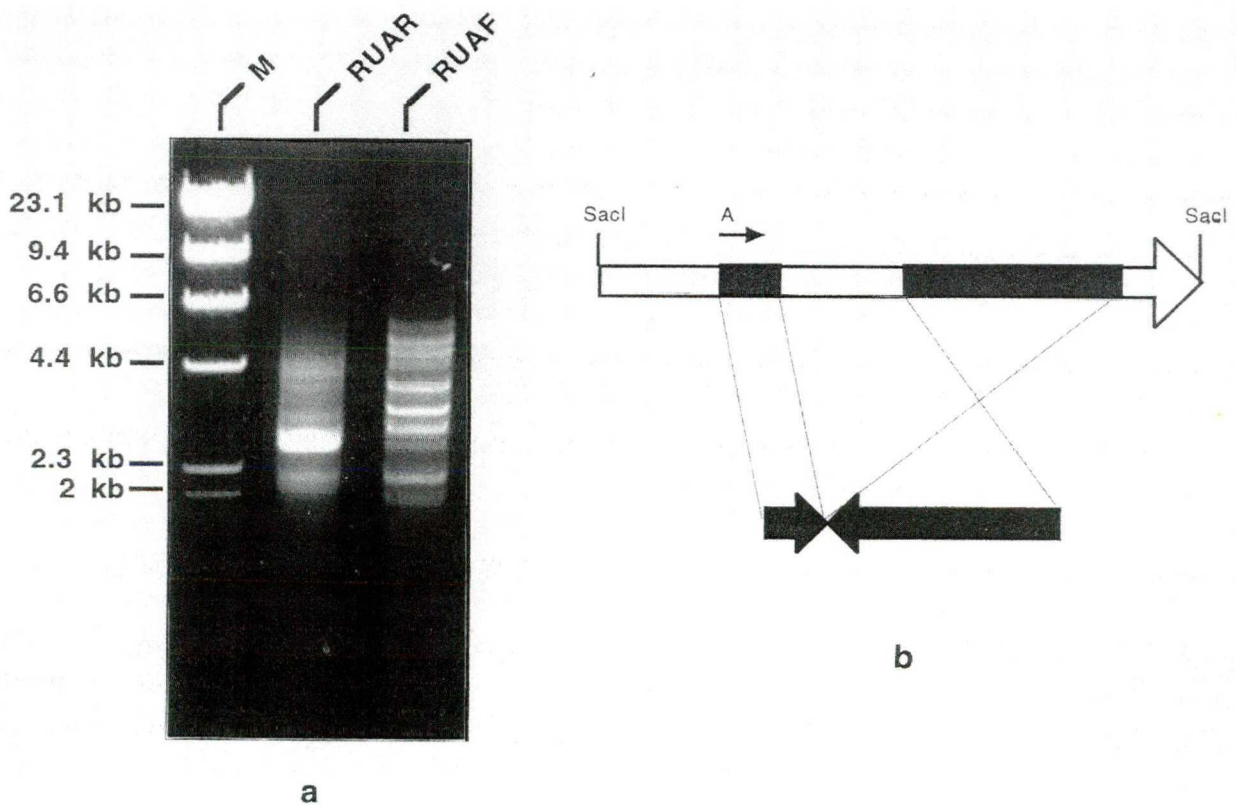


felhasználásával. Az egy primeres PCR amplifikációk eredményét mutatja a 10a. ábra. A repetitív egység 3' végére (RUAR: unit "a", 2382-2415) tervezett primer használata egy erős ~2.7 kb-os és néhány kevésbé intenzív fragment amplifikációjához vezetett. Az egység 5' végére (RUAF: unit "a", 800-833) specifikus primer ugyanakkor több, hasonló intenzitású fragment amplifikációját eredményezte. A fordított ismétlődő szakaszok létét szekvencia analízissel is megerősítettük. Az egyik 982 bp méretű klónozott PCR fragment szerkezetét a 10b. ábra mutatja. Az amplifikációhoz használt primerrel homológ szekvencia részlet a fragment egyik végéről hiányzik. Ez az eredmény nem meglepő ugyanis jól ismert, hogy fordított ismétlődéseket tartalmazó plazmid molekulák *E. coliban* nagyfokú instabilitást mutatnak.

A szatelliteszerű DNS elemek amplifikációjában, az uralkodó nézet szerint a testvér kromatidák között lejátszódó egyenlőtlen *crossing overeknek* tulajdonítanak döntő szerepet. Habár ez a mechanizmus jól megmagyarázhatja a pericentrikus szatellit szakaszok méretében megfigyelt, sokszor egyedekre jellemző nagymértékű polimorfizmust (96), nem ad magyarázatot a fordított ismétlődések kialakulására. Tenyésztett emlős sejtekben, növekvő drog koncentrációk alkalmazásával általában könnyen előidézhethetjük a szelektált gén (és sokszor annak kiterjedt környezetének) a nagyfokú amplifikációját (43). Az ilyenkor megfigyelt amplifikált struktúrákban az esetek többségében fordított ismétlődések dominálnak (97). Ez a megfigyelés látszólagos ellentmondásban van a szatellitek esetében uralkodó tandem szerveződéssel. Feloldhatja ezt az ellentmondást az a feltételezés, hogy a fordított ismétlődéseket eredményező mechanizmusok a DNS szakaszok amplifikációjának korai stádiumában dominálnak, majd később a tandem struktúrákat kialakító mechanizmusok kerülnek túlsúlyba. Ez a feltételezés maga után vonja, hogy a szatellitekkel rokon, fordított ismétlődő szakaszok a szatellit kialakulásának kezdeti időszakában jöttek létre. Mivel feltételezhetően a fordított struktúrák topológiailag megnehezítik a *crossing overen* alapuló homogenizációs folyamatok lejátszódását az ilyen szakaszokon a mutációk akkumulálódását



várhatjuk. Ezzel szemben a vizsgált, HC2 szatellittel rokon, fordított ismétlődést tartalmazó DNS szakasz több mint 97% homológiát mutatott az "a" típusú repetitív egység megfelelő szakaszaival. Ez arra utal, hogy a fordított struktúra kialakulása evolúciósan új keletűnek tekinthető. Valószínű tehát, hogy azok az amplifikációs mechanizmusok amelyek fordított ismétlődések kialakulásához vezetnek a szatellitok evolúciójának későbbi szakaszaiban is aktívak maradnak, feltételezhetően azonban a centromer közeli heterokromatikus régiókban kevésbé aktívak, szemben az eukromatikus kromoszóma szakaszokkal, ahol a drogrezisztenciák kialakulásáért felelős fő amplifikációs mechanizmusok lehetnek.



**10. ábra.** A HC2 szatellit fordított ismétlődéseinek vizsgálata egy primeres PCR amplifikációkkal (a). Az egyik klónozott és szekvenált fordított ismétlődő szakasz szerkezete (b).

## 5. ÖSSZEFOGLALÁS

Az alacsonyabb rendű eukarióták centromerikus DNS szakaszainak azonosításában és funkcionális jellemzésében elért jelentős sikerek ellenére az emlős centromer azonosítása még mindig várat magára. Nagyban hozzájárul ehhez a sikertelenséghez a pericentrikus kromoszóma régiók különleges DNS összetétele. Magasabbrendű eukariótákban a centromerikus kromoszóma szakaszok túlnyomórészt tandem szerveződő, sokszor faj vagy kromoszóma specifikus repetitív DNS elemekből ún. szatellitekből épülnek fel. A szatellitek pericentrikus koncentrálódása több szempontból is megnehezíti a funkcionális emlős centromer definiálását. A rövid repetitív egységekből álló szatellit szekvenciák általában nagy mértékű instabilitást mutatnak konvencionális vektorokba klónozva. Emiatt a hagyományos génbankokban a centromerikus kromoszóma szakaszok DNS elemei alulreprezentáltak. Ez a tény, valamint az emlős centromerikus DNS konzervativitásának látszólagos hiánya felveti annak az igényét, hogy megkíséreljük centromerikus kromatin frakciók izolálását, majd ezek felhasználásával centromer specifikus DNS bankok létrehozását.

Centromerikus kromatin frakciók izolálására kidolgozott módszerekről több közlemény is beszámol. Az eljárások többségükben a kromoszóma karok nukleáz kezelés révén történő szelektív eltávolítására épülnek (98, 99, 100). Ennek az a megfigyelés szolgál alapjául miszerint az emlős kromoszómák pericentrikus régiói kompakt, nukleázokkal szemben viszonylag ellenálló kromatin struktúrát vesznek fel. Egy másik beszámolóban a kromoszómák ultrahangos fragmentálását követően, centromer fehérjéket felismerő ellenanyagokkal végrehajtott immunoprecipitáció révén kísérelték meg centromerikus kromatin izolálását (101). Habár az ismertett módszerek eredményezhetik a centromerikus kromatin frakció szelektív dúsulását, mindkét megközelítésnek lehetnek hátrányai. A nukleázokat alkalmazó eljárások arra az alapfeltevésre épülnek, hogy a centromerikus



kromatin nagy mértékben védett DN-áz emésztéssel szemben. Ez a feltételezés azokra a citológiai megfigyelésekre alapul, amelyek a centromerikus kromoszóma régiók intakt megőrződését mutatják korlátozott DN-áz kezelés során. Nincsenek viszont adatok arra nézve, hogy a centromer teljes DNS tartalma megőrződik-e ilyen viszonyok mellett. Lehetséges, hogy a nukleáz kezeléssel előállított kromatin frakciók nem reprezentálják a centromer teljes DNS tartalmát. Az ultrahang illetve egyéb mechanikai behatások alkalmazása ugyancsak aggályos lehet. Az ilyen módszerekkel végrehajtott kromoszóma fragmentáció nehezen szabályozható. Elégtelen mértékű fragmentáció nem vezethet a centromerikus kromatin megfelelő fokú dúsulásához, ugyanakkor túlzott mechanikai behatások a centromer-kinetochor komplex sérüléséhez vezethetnek, ami viszont újra csak bizonyos centromerikus DNS frakciók elvesztését eredményezheti.

A fentebb részletezett problémákra jelenthet megoldást az általunk alkalmazott *in vivo* kromoszóma fragmentáció módszere. Korábbi megfigyelések bizonyítják, hogy a hidroxürea-koffein kezeléssel indukált korai kromatin kondenzáció (*premature chromatin condensation: PCC*) során a kromatin többségétől elváló centromer-kinetochor komplex mind struktúráját mind funkcióját megőrzi (91). Ezeknek a centromerikus kromoszóma daraboknak az immunoprecipitációja jó eséllyel szolgáltathat olyan kromatin frakciót amelyben a teljes, funkcionális centromer-kinetochor komplex DNS tartalma reprezentálva van. Az itt ismertetett kísérletek során kínai hörcsög sejtekben *in vivo* kromoszóma fragmentációt indukáltunk. Ezt követően anti-centromer ellenanyagokat tartalmazó humán autoimmun szérum felhasználásával immunoprecipitáltuk a centromerikus kromoszóma darabokat. Az így módon kapott kromatin frakcióból DNS-t tisztítottunk és azt plazmid vektorba klónoztuk. Nyolc plazmid klónt véletlenszerűen kiválasztottunk és a klónozott DNS fragmentek kromoszómális lokalizációját *in situ* hibridizációval vizsgáltuk metafázisos hörcsög kromoszómákon. Az egyik 536 bp hosszú DNS fragment (HUCAFF170) erős hibridizációs jelet mutatott a 2. kínai hörcsög kromoszóma pericentrikus régiójában. Az



általunk itt alkalmazott módszerhez hasonló eljárást mások is sikeresen használtak (80). Ouspenski és munkatársa az 1. kínai hörcsög kromoszóma centromerikus régiójából izoláltak egy 6 kb méretű, egy példányban előforduló DNS szakaszt. Ők három, véletlenszerűen kiválasztott klón vizsgálata során azonosították az új centromerikus szekvenciát. Összesen tehát 11 klón analízise két új centromerikus DNS elem izolálásához vezetett. Ez az adat az itt részletezett módszer potenciális alkalmazhatóságát mutatja. Az analizált alacsony klón szám azonban még nem teszi lehetővé, hogy megbízhatóan megítéljük a centromerikus kromatin dúsulásának mértéket. Ennek egyértelmű eldöntése további vizsgálatokat igényel.

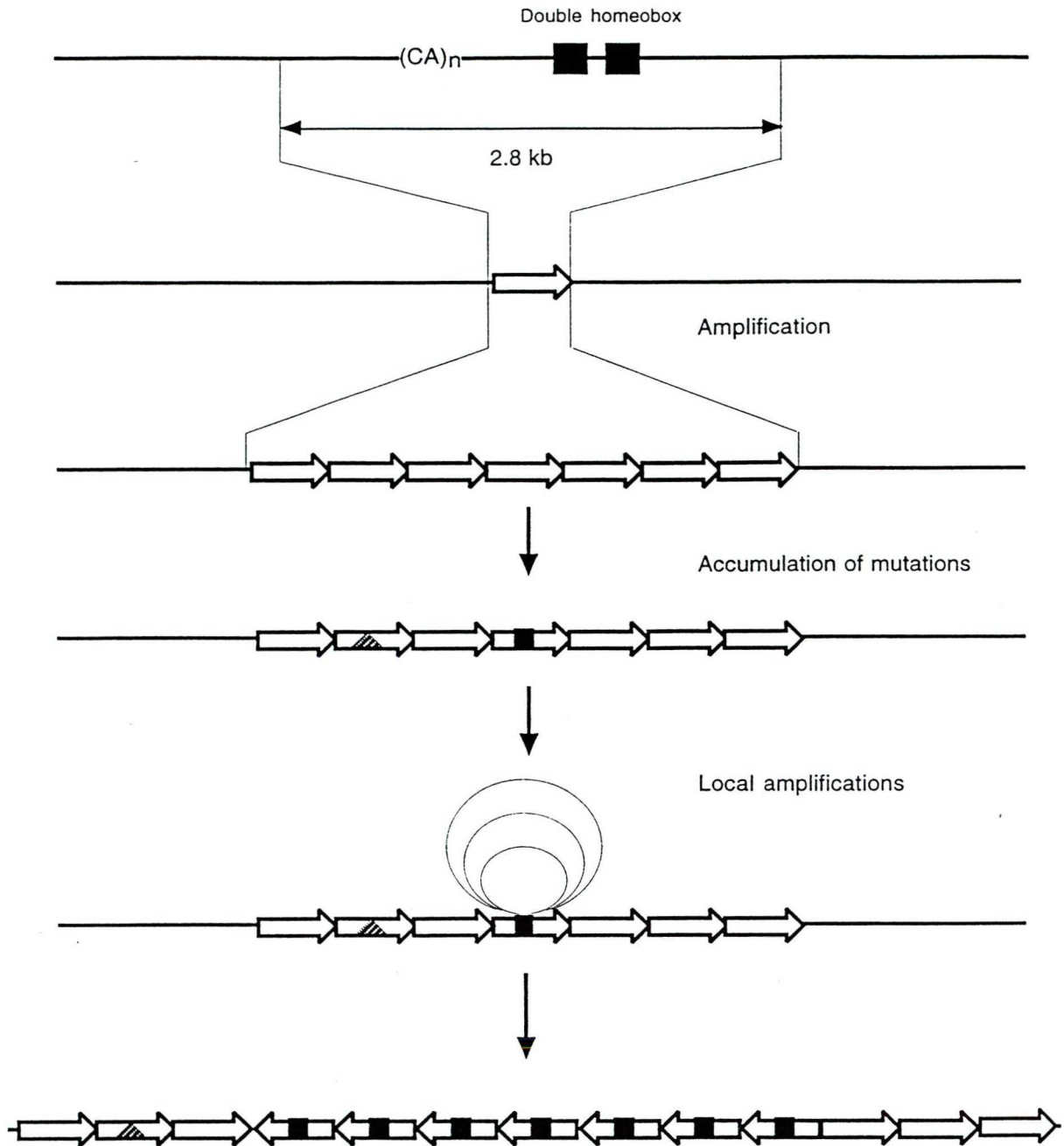
A továbbiakban megvizsgáltuk az azonosított új centromer szekvencia genomikus szerveződését. A HUCAFF170 klón próbaként való használatával izolált 13.3 kb méretű lambda fág klón (HC90) szerkezetének elemzése kimutatta, hogy az eredeteileg azonosított 536 bp-os DNS fragment homológiát mutat, egy tandem szerveződő repetitív DNS elem 2.8 kb méretű ismétlődő egységeivel. Genomikus *Southern* hibridizációk megerősítették, hogy az új szatellit-szerű DNS szekvencia a 2. kínai hörcsög kromoszóma pericentrikus régiójának egyik fő alkotó eleme (HC2 szatellit).

A 2.8 kb méretű *SacI* egységek két rövid szakaszon homológiát mutatnak a homeotikus gének DNS-t kötő homeobox régióival. Ezekről a szakaszokról eltekintve a repetitív egységek nem mutatnak homológiát más ismert DNS elemmel. Az analizált *SacI* egységek határán azonosítható 1062 bp hosszú nyílt leolvasási keret egy olyan 353 aminosavból álló potenciális polipeptidet kódolhat amely két, a homeoboxsal magas szintű homológiát mutató peptid részletet tartalmaz. Habár a nyílt leolvasási keret előtt promóterek konszenzus szekvenciával homológiát mutató szakaszok nem azonosíthatók, nem zárhatjuk ki, hogy néhány, a repetitív egységbe ágyazott ORF-el rokon DNS szakasz esetenként expresszálandó. Ezt a feltevést megerősíteni látszik az RT-PCR kísérletek eredménye is. Az általunk azonosított, homeoboxsal rokon repetitív szekvenciákat főemlősökből is leírtak. A humán 4. kromoszómán elhelyezkedő 3.3 kb-os repetitív egységgel rendelkező D4Z4

elemet egy domináns öröklés menetet mutató betegség (FSHD) kapcsán írták le (24). A kiterjedt kutatások ellenére még ma sem tisztázott a D4Z4 repetitív egységeinek deléciója és az FSHD manifesztációja közötti kapcsolat, habár a D4Z4 elemek heterokromatikus lokalizációja pozíció effektus variegációra utalhat. A legújabb vizsgálatok kimutatták, hogy a D4Z4 elemmel rokon szekvenciák főemlősökben is megtalálhatók (95). Ez a megfigyelés valamint az általunk hörcsögben azonosított új repetitív DNS elem felveti annak a lehetőségét, hogy a homeoboxsal rokon repetitív DNS elemeknek valamilyen ma még nem ismert evolúciósan konzervált szerepe lehet.

A HC90 fág klónban azonosított 2.8 kb-os SacI egységek szekvenciáinak összehasonlítása az ismétlődő egységek különböző szakaszainak eltérő konzervativitását mutatták ki. Míg az "a" és "b" típusú SacI egységek végei majdnem azonosak addig a középső 1.5 kb-nyi szakaszukon csak 88.5% homológiát mutatnak egymással. A genomikus *Southern* hibridizációk ugyancsak a repetitív egységek középső szakaszainak nagyfokú heterogenitására utaltak. Érdeemes megjegyezni, hogy a repetitív egységek nagy mértékben konzervált szakaszai egybeesnek a kettős homeoboxot tartalmazó ORF elhelyezkedésével. Ez felveti annak a lehetőségét, hogy ezen szakaszok relatív konzervativitása valamilyen funkcióval kapcsolatos szelekciós nyomásra vezethető vissza. Nem zárhatjuk ki azonban annak a lehetőségét sem, hogy a középső szakaszon koncentrálódó különleges szekvenciái elemek ("egyszerű" szekvenciák, hosszú A-T szakaszok) okozzák ennek a régióknak a fokozott instabilitását.

A *pulsed field* gélelektroforézissel kombinált *Southern* hibridizáció a HC2 szatellit nagyméretű homogén blokkjait mutatta ki, míg az egy primerrel végrehajtott PCR kísérletek fordított ismétlődések létét bizonyították. Összevetve ezeket az adatokat felvázolhatjuk a HC2 szatellit kialakulásához vezető legvalószínűbb evolúciós útvonalat (11. ábra).



11. ábra A HC2 szatellit kialakulásához vezető lehetséges evolúciós útvonal vázlata.

A szatellit őseül feltételezhetően egy homeotikus génekkel homológiát mutató DNS szakasz szolgálhatott. Ebben a pszeudogén-szerű szekvenciában először a homeoboxot meghatározó szekvencia részlet duplikációja játszódhatott le. A DNS elem nagyobb fokú



amplifikációjának valószínűleg feltétele volt egy rekombinációs "hot spot"-ként működő  $(CA)_n$  szakasznak a pszeudogén közelébe történő beépülése. Az elsődleges amplifikációt további, lokális szaltatorikus amplifikációk követhették, amelyek a kiterjedt homogén blokkok kialakulását eredményezték. A szatellit amplifikációja különféle mechanizmusok révén történhetett. A testvér kromatidák között lejátszódó egyenlőtlen *crossing over* mellett egyéb, fordított ismétlődéseket kialakító amplifikációs folyamatok is lejátszódtak. A lokális amplifikációk egy vagy néhány repetitív egységet magukba foglaló "rolling circle" struktúrák révén történhettek. Ugyanakkor a már kialakult nagyméretű homogén blokkok egy lépésben lejátszódó duplikációja is bekövetkezhetett amint azt az azonos méretű XbaI domének léte valószínűvé teszi. A HC2 szatellit repetitív egységeinek szintjén lejátszódó amplifikációs események mellett a repetitív egységen belüli "egyszerű" szekvenciák párhuzamos evolúciójának is szemtanúi lehetünk. A szatellit ismétlődő egységének őseül szolgáló DNS elem tartalmazhatott egy kitüntetett helyzetű tetranukleotid szakaszt, amely a 2.8 kb-os egységek amplifikációját követően, az egyes repetitív egységekben egymástól függetlenül vezethetett a tetranukleotid elemekből felépülő "egyszerű" szekvenciák kialakulásához. Összefoglalva tehát a HC2 szatellit mai formáját hierarchikusan egymásra épülő szinteken lejátszódó, különböző mechanizmusokat használó amplifikációs folyamatok alakították ki.

A szatellitek sokszor faj sőt extrém esetekben kromoszóma specifikusak is lehetnek. Az általunk azonosított HC2 szatellit sem mutat szekvencia homológiát más fajok szatellit elemeivel. *In situ* hibridizáció ugyanakkor azt is kimutatta, hogy az új szatellit a kínai hörcsög kromoszómák közül is csak a 2. kromoszómán található meg. Ezt megelőzően kínai hörcsögből csak az emlősök telomerikus DNS szakaszaival homológ,  $(TTAGGG)_n$  összetételű repetitív elem volt ismert (102). Ez a kromoszómák Robertsoni fúziójának maradványaként tekinthető szekvencia elem az 1. és 2. kromoszómák kivételével az összes hörcsög kromoszóma pericentrikus régiójában megtalálható. A telomerrel homológ pericentrikus DNS elemek hiánya a két nagy szubmetacentrikus kromoszómának a többi

kromoszómától eltérő evolúciós eredetére utal. Az általunk azonosított 2. kromoszóma specifikus szatellit ezt a feltételezést megerősíti.

A szatellitek funkciója még ma sem egyértelműen eldöntött kérdés. A jelenleg rendelkezésünkre álló adatok a pericentrikus szatelliteknek a testvér kromatidák összetartásában betöltött szerepével vannak leginkább összhangban. Specifikus funkciók és a szatellitek közötti kapcsolat megállapítását nagy mértékben megnehezíti a szekvencia konzervativitás látszólagos hiánya. Ezt az ellentmondást oldhatja az a feltételezés, hogy a pericentrikus heterokromatin sajátosságainak kialakulásában a szatellitek ismétlődő egységeinek szekvenciája csak másodlagos szerepet tölt be és a centromer funkciójához esszenciális kromatin konformációt lényegileg bármilyen, szabályos periodicitást mutató DNS elem képes kialakítani. Ezt a feltételezést erősítik meg azok az újabb vizsgálatok amelyek tandem ismétlődő transzgenek heterokromatin-szerű sajátosságait mutatták ki *Drosophilában* és növényekben (25, 26). Nem zárhatjuk ki azonban annak a lehetőségét sem, hogy a pericentrikus szatellitekben általánosan előfordulnak rövid konzervált, potenciális fehérjékötő helyként szolgáló szekvencia elemek. A szatellitek ismétlődő egységeinek többségét kitevő "töltelék" DNS szakaszok funkciója ez esetben az lehet, hogy a konzervált elemeket megfelelő elrendezésben és távolságban tartsa. Az ilyen jellegű térkitöltő funkció csak kis mértékű szelekciós nyomást fejthet ki ezen DNS szakaszok szekvenciájának megőrződésére.

Egyes feltételezések szerint a szatellit szekvenciák az interfázisos sejtmag szerveződésében is fontos funkciót tölthetnek be (38). Könnyen beláthatjuk, hogy az általunk azonosított szatellitekhez hasonló, változatos, esetenként kromoszóma specifikus szatellitek részt vehetnek a sejtmagi rend kialakításában elősegítve például azt, hogy egyfajta "cimkeként" szolgálva a homológ kromoszómák interfázisban is felismerjék egymást. A kromoszómák ilyen fajta megjelölése fontos lehet abban, hogy adott kromoszómák az általuk hordozott gének megfelelő expresszióját biztosító sejtmagi kompartmenteket foglalják el.



## 6. SUMMARY:

One of the most interesting distinctive features of the eukaryotic genome is that it contains a large amount of seemingly useless repetitive DNA elements. The first data indicating the existence of such repetitive DNA fractions obtained from experiments studying the re-association kinetics of eukaryotic genomic DNA. Further studies elucidated that the repetitive DNA elements can be classified into two major groups according to their organization. The interspersed repetitive elements scattered more or less evenly throughout genome. This phylogenetically diverse group includes DNA elements, which at least at one point of their evolution were mobile. Recent studies revealed that certain members of this group function as telomere in Dipteras, while others may play a role in the regulation of eukaryotic DNA replication. Despite the recent advances in our understanding of the organization of the eukaryotic genome, the role of the majority of the interspersed repetitive DNA elements is still enigmatic.

The other major class of the repetitive DNA sequences is characterized by DNA segments (repeat units) reiterated up to several thousand times in a tandem manner. The size of the repeat unit can vary from a few to several thousands basepairs. One type of the tandem repetitive sequences the so-called satellites are usually localized in the centromeric regions of the chromosomes. Although, DNA sequences capable of forming an active centromere in mammalian cells have not been identified yet, many observations indicate that the satellite sequences perform at least part of this function.

There are several obstacles that hinder the identification of the minimal mammalian centromere DNA. One of these difficulties is associated with the fact that satellite DNA sequences are usually difficult to handle by standard molecular techniques. The other, even more serious deficiency is that there is no reliable functional assay to identify the minimal mammalian centromere DNA. One possible solution to these problems is to isolate the



chromosome regions corresponding to the functional centromere and study the DNA components of such chromosome segments. Earlier several procedures have been developed for isolation of chromatin fractions enriched in centromeric DNA. These techniques usually rely on the observation that centromeric regions of the mitotic chromosomes are relatively resistant to nuclease digestion. Although several centromeric DNA elements have been isolated by using these methods, it is not clear whether all of the DNA components of the centromere are preserved under these conditions. It is possible that the sequence composition of DNA obtained from such chromatin fractions may not be representative of the entire centromere.

In our studies we used a novel approach to enrich the centromeric chromatin fraction of Chinese hamster chromosomes. We induced *in vivo* chromosome fragmentation in cultured hamster cells by hydroxyurea-caffeine treatment. It has been demonstrated earlier that the centromeres detached by this treatment from the rest of the chromosomes retain their mitotic functions. Following the hydroxyurea-caffeine treatment we isolated the chromosome fragments from the cells and immunoprecipitated the centromere-kinetochore complexes by an autoanti-serum that reacted with centromere proteins. By screening the plasmid library, constructed from the immunoprecipitated centromeric chromatin, we identified a DNA fragment (HUCAFF170) that we were able to localize to the centromeric region of the 2. Chinese hamster chromosome by *in situ* hybridization. Using this DNA fragment as a probe we isolated a lambda phage clone (HC90) from a hamster DNA library carrying a 13.3 kb DNA insert. The restriction mapping of this phage clone revealed that the HUCAFF170 fragment is part of the 2.8 kb long repeat unit of a novel tandemly organized, satellite-like repetitive sequence. This DNA element is the only known component of the centromeric region of the 2. Chinese hamster chromosome. Copy number estimation indicated that 2,500-5,000 copies of this repeat unit are present on each hamster chromosome 2., therefore the novel satellite is one of the major components of the centromeric region of these

chromosomes (HC2 satellite). We performed the sequence analysis of two such repeat units. The sequenced repeat units exhibited a high degree of homology to each other. On the other hand no similarity was observed to other known DNA sequences, except for two short segments that showed a significant homology to the highly conserved homeobox region of the homeotic genes. Interestingly, these DNA segments are part of a 1062 bp long open reading frame. RT-PCR data indicated that some of these ORFs might be expressed in Chinese hamster cells. Recently similarly organized, homeobox containing repetitive DNA elements have been described from human, primates and old world monkeys. The human D4Z4 element has also been implicated in an autosomal recessive disorder. Although the novel hamster repetitive element does not show homology to these sequences aside from the homeobox region, it is interesting to speculate whether the tandemly organized, homeobox containing repetitive DNA elements may have some important, evolutionary conserved function in these organisms.

In the central part of the repeat units we saw concentration of different simple repeats and long A-T rich tracts. It is possible that these sequence elements may contribute to the observed heterogeneity of these regions. The long  $(CA)_n$  tracts might also have been an important role in the evolution of the novel repetitive element by acting as a recombination hot spot.

Although genomic Southern hybridizations indicated that the majority of the repeat units organized in a tandem manner we could also detect related inverted sequence elements. These inverted structures may be a result of different amplification processes. By pulsed field gel electrophoresis we could also detect large homogeneous domains of the tandem array. Most likely these homogeneous domains generated by local saltatory amplification events. Additionally, our results also indicate that large-scale duplications of long segments of the tandem array also took place. In conclusion, these data suggest that the amplification events



that led to the formation of the chromosome 2. satellite array took place on different scales and involved different amplification mechanisms.

The function of the satellite sequences in higher eukaryotes if any, is still disputed. Although up to now lots of data accumulated that argue for their role in the chromosome segregation, the apparent lack of sequence conservation between centromeric satellites of different species makes it difficult to attribute specific functions to them. The heterogeneity of the centromeric DNA sequences is in contradiction with the observation that in many cases the chromosomes of both parental cell lines are stable in inter-species cell hybrids, indicating that at some degree the centromeres of the two species must be conserved. A possible solution to this paradox was provided by recent observations that common sequence motifs can be found in otherwise highly diverged centromeric satellites of different species. It is suggested that these sequence motifs can bind conserved centromere proteins. If these short sequence motifs are indeed the only functionally important elements of the satellites it is a reasonable assumption that the selection pressure to rest of the repeat units is much lower, leading to a much higher diversity of these regions. A different set of data also indicates that the lack of sequence conservation does not necessarily exclude similar functions. It has been shown that tandemly repeated sequence elements (transgenes) can adopt heterochromatin-like chromatin conformation. In light of these discoveries, we can speculate that although the novel satellite sequence described here does not show extensive homology to other centromeric satellites, it may contain conserved binding sites of centromeric proteins, while the periodical repetition of its repeat unit can contribute to the formation of the characteristic chromatin conformation that may be essential for the proper centromere function.



## 7. IRODALOM:

- 1., Singer, M.F. (1982) *International Review of Cytology* 76, 67-112
- 2., Smit, A. F.A. (1996) *Current Opinion in Genetics & Development* 6, 743-748
- 3., Smit, A. F.A. (1995) *Nucl. Acids Res.* 23, 98-102
- 4., Herrera, R.J. & Wang, J. (1991) *Genetica* 84, 31-37
- 5., Szemraj, J. et al (1995) *Gene* 152, 261-264
- 6., Boeke, J.D. (1997) *Nature Genetics* 16, 6-7
- 7., Sassaman, D.M. et al (1997) *Nature Genetics* 16, 37-43
- 8., Morse, B. et al (1988) *Nature* 333, 87-90
- 9., Kazazian, H.H. et al (1988) *Nature* 332, 164-166
- 10., Doolittle R.F. et al (1989) *Q. Rev. Biol.* 64, 1-30
- 11., Smit, A. F.A. (1993) *Nucl. Acids Res.* 21, 1863-1872
- 12., Smit, A. F.A. & Riggs, A. D. (1996) *Proc. Natl. Acad. Sci. USA* 93, 1443-1448
- 13., Capy, P. et al (1994) *Genetica* 93, 161-170
- 14., Fry, K. & Salser, W. (1977) *Cell* 12, 1069-1084
- 15., Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517
- 16., Delidakis, C. et al (1989) *Current Opinion in Cell Biology* 1, 488-496
- 17., Warburton P.E. & Willard H. F. (1990) *J. Mol. Biol.* 216, 3-16
- 18., Strand, M. et al (1993) *Nature* 356, 274-276
- 19., Dover, G. (1982) *Nature* 299, 111-117
- 20., Pavelitz, T. et al (1995) *EMBO J.* 14, 169-177
- 21., Pardue, M.L. & Gall, J.G. (1970) *Science* 168, 1356-1358
- 22., Lohe, A.R. & Hilliker, A.J. (1995) *Current Opinion in Genetics & Development* 5, 746-755
- 23., Allshire, R.C. et al (1994) *Cell* 76, 157-169

- 24., Lyle, R. et al (1995) *Genomics* 28, 389-397
- 25., Dorer, D.R. & Henikoff, S. (1994) *Cell* 77, 993-1002
- 26., Ye, F. & Signer, E.R. (1996) *Proc. Natl. Acad. Sci. USA* 93, 10881-10886
- 27., Eberl, D.F. et al (1993) *Genetics* 134, 277-292
- 28., Bayne, R.A. et al (1994) *Hum. Mol. Genet.* 3, 539-546
- 29., Heller, R. et al (1996) *Proc. Natl. Acad. Sci. USA* 93, 7125-7130
- 30., Vogt, P. (1990) *Hum. Genet.* 84, 301-336
- 31., Hamada, H. et al (1984) *Mol. Cell. Biol.* 4, 2610-2621
- 32., Herbert, A.G. et al (1993) *Proc. Natl. Acad. Sci. USA* 90, 3339-3342
- 33., Tautz, D. & Schlotterer, C. (1994) *Current Opinion in Genetics & Development* 4, 832-837
- 34., Meyne et al (1989) *Proc. Natl. Acad. Sci. USA* 86, 7049-53
- 35., Tyler-Smith, C. & Willard, H.F. (1993) *Current Opinion in Genetics & Development* 3, 390-397
- 36., Jeffreys A.J. et al (1985) *Nature* 316, 76-79
- 37., Earnshaw, W.C. & Pluta, A.F. (1994) *BioEssays* 16, 639-643
- 38., Manuelidis, L. (1990) *Science* 250, 1533-1540
- 39., Monaco, A.P. & Larin, Z. (1994) *TIBTECH* 12, 280-286
- 40., Huxley, C. (1994) *Gene Therapy* 1, 7-12
- 41., Deshpande, A.M. & Newlon, C.S. (1992) *Mol. Cell. Biol.* 12, 4305-4313
- 42., Stillman, B. (1989) *Annu. Rev. Cell. Biol.* 5, 197-245
- 43., Hamlin, J.L. et al (1991) *Prog. Nucleic Acid Res. Mol. Biol.* 41, 203-239
- 44., Vaughn, J.P. et al (1990) *Cell* 61, 1075-1087
- 45., Caddle, M.S. et al (1990) *J. Mol. Biol.* 211, 19-33
- 46., Watson, J.D. (1972) *Nature N. Biol.* 239, 197-201
- 47., Pardue, M.L. (1994) *Current Opinion in Genetics & Development* 4, 845-850

- 48., Biessmann, H. & Mason J.M. (1994) *Chromosoma* 103, 154-161
- 49., Blackburn, E.H. (1994) *Cell* 77, 621-623
- 50., Williamson, J.R. et al (1989) *Cell* 59, 871-880
- 51., Morin, G.B. (1989) *Cell* 59, 521-529
- 52., Greider, C.W. (1990) *Bioessays* 12, 363-369
- 53., Kim, N.W. et al (1994) *Science* 266, 2011-2015
- 54., Palladino, F. & Gasser, S.M. (1994) *Current Opinion in Genetics & Development* 6, 373-379
- 55., Danilevskaya, O. et al (1994) *Chromosoma* 103, 215-224
- 56., Fang-miin, S. & Levis, R.W. (1994) *Proc. Natl. Acad. Sci. USA* 91, 12510-12514
- 57., Rattner, J.M. (1991) *BioEssays* 13, 51-56
- 58., Yen, T.J. et al (1992) *Nature* 359, 536-539
- 59., Pluta, A.F. et al (1995) *Science* 270, 1591-1594
- 60., Cooke, C.A. et al (1987) *J. Cell. Biol.* 105, 2053-2067
- 61., Hadlaczky, Gy. et al (1986) *Exp. Cell. Res.* 167, 1-15
- 62., Sullivan, B.a. et al (1996) *Environ. Mol. Mutagen.* 28, 182-191
- 63., Sunkel, E.S. & Coelho, P.A. (1995) *Current Opinion in Genetics & Development* 5, 756-767
- 64., Clarke, L. & Carbon, J. (1980) *Nature* 287, 504-509
- 65., Hegemann, J.H. & Fleig, U.N. (1993) *BioEssays* 15, 451-460
- 66., Clarke, L. et al (1993) *Cold Spring Harbor Symp. Quant. Biol.* 58, 669-676
- 67., Baum, M. et al (1994) *Mol. Biol. Cell.* 5, 747-761
- 68., Marschall, L. & Clark, L. (1995) *J. Cell. Biol.* 128, 445-454
- 69., Murphy, T.D. & Karpen, G.H. (1995) *Cell* 82, 599-609
- 70., Willard, H.F. (1990) *Trends Genet.* 6, 410-416
- 71., Vig, B.K. (1994) *Mutat. Res.* 309, 1-10



- 72., Grady, D.L. et al (1992) *Proc. Natl. Acad. Sci. USA* 89, 1695-1699
- 73., Abad, J.P. et al (1992) *Proc. Natl. Acad. Sci. USA* 89, 4663-4667
- 74., Willard, H.F. & Wayne, J.S. (1987) *Trends Genet.* 3, 192-198
- 75., Masumoto, H. et al (1989a) *J. Cell. Biol.* 109, 1963-1973
- 76., Masumoto, H. et al (1989b) *Exp. Cell. Res.* 181, 181-196
- 77., Kipling, D. et al (1995) *Mol. Cell. Biol.* 15, 4009-4020
- 78., Kipling, D. et al (1994) *Chromosoma* 103, 46-55
- 79., Cooper, K.F. et al (1992) *J. Mol. Biol.* 228, 421-432
- 80., Ouspenski, I.I. & Brinkley, B.R. (1993) *J. Cell. Sci.* 105, 359-367
- 81., Kipling, D. (1997) *Trens Genet.* 13, 141-145
- 82., Earle, E. et al (1992) *Cytogenet. Cell Genet.* 61, 78-80
- 83., Voullaire, L.E. et al (1993) *Am. J. Hum. Genet.* 52, 1153-1163
- 84., Haaf, T. et al (1992) *Cell* 70, 681-696
- 85., Larin, Z. et al (1994) *Hum. Mol. Genet.* 3, 689-695
- 86., Harrington, J.J. et al (1997) *Nature Genetics* 15, 345-355
- 87., Hadlaczky, Gy. et al (1989) *Chromosoma* 97, 282-288
- 88., Sambrook, J. et al (1989) *Molecular Cloning: A Laboratory Manual*, CSH
- 89., Altschul, S.F. et al (1990) *J. Mol. Biol.* 215, 403-410
- 90., Schlegel, R. & Pardee, A.B. (1986) *Science* 232, 1264-1266
- 91., Brinkley, B. R. et al (1988) *Nature* 336, 251-254
- 92., Skowronski, J. et al (1984) *J. Mol. Biol.* 177, 399-416
- 93., Martinez-Balbas, A. et al (1990) *Biochemistry* 29, 2342-2348
- 94., Hyrien, O. et al (1987) *EMBO J.* 6, 2401-2408
- 95., Clark, L.N. et al (1996) *Chromosoma* 105, 180-189
- 96., Marcais, B. et al (1991) *Hum. Genet.* 86, 311-316
- 97., Stark, G.R. et al (1989) *Cell* 57, 901-908

98., Rattner, J.B. et al (1978) *Chromosoma* 66, 259-268

99., Lica, L. & Hamkalo, B. (1983) *Chromosoma* 88, 42-49

100., Valdivia, M.M. & Brinkley, B.R. (1985) *J. Cell. Biol.* 101, 1124-1134

101., Hadlaczky, Gy. et al (1991) *Proc. Natl. Acad. Sci. USA* 88, 8106-8110

102., Meyne, J. et al (1990) *Chromosoma* 99, 3-10

## 8. MELLÉKLETEK

## 1. Melléklet: Az "a" típusú ismétlődő egység nukleotid szekvenciája:

1 GAGCTCGCCC AAAAGACCTA TTTACCGGAA GACACGATTG  
 41 ACATATGGTT TCAGAACAGA AGAGCACGCC ACCGTGCCGC  
 81 CTGCACGGGT AGAGTTCCGA GTCCAGATTG CCTTTCCTCG  
 121 CCTGAGCCCA ATGAGGCAGT GCACGGCAGC GAGACTGAGC  
 161 GTGCACAGGA CAGCTTGTG CCACTGGCTG CTGCATCACC  
 201 TGTGGGCACA GGCACTTTGA CCATGACTTG CCCATCAACC  
 241 TCCTACCGAA TGTCTGGCC TGCAGAGACG GCACAGCCCT  
 281 CTCCACCAGA CGTCCCAGAG ACACACACCG TTGCATGGAA  
 321 CACCTGCGCT GGGCAGCTCT TTCTTGATGC AGTGATGAAA  
 361 GAGATCCAAG ATGACCAAGA TCTCCGACCC CTGCAGGACT  
 401 CTCTGGATCC CTACGATCAA TGGGAGCAGA TGCAGCCAAC  
 441 CCTTGAGCCC CCTCTGGACG AAGAAGTCTA TCAAGCTCTA  
 481 CTGGATGACC TCTGTGACAC ACACTGCAAG GCTTCATCGT  
 521 CTCATCACGG ATGCCGTCTT CTTCCAGAAA GCAAGACCCCT  
 561 TACACCAAAC CCCAACACAG GACAAAATGG ACCTCCAACG  
 601 CCCCTGCCCC TGTTCGGAG CCCATGGACA GGAATCCCCT  
 641 GACCCTAAGC AACTGGGAAA AATCACCAGA AACAGGAAGT  
 681 CGGGTGTGGA AGGCCATCCC AGGAGGCCGA GGACTIONAAA  
 721 ATTCCCTGGC AGGAGGCGTT TGCTGGAGAC TAGAGGGAAA  
 761 TTTCCCCAGT CATAACACAGG AAGACATCTC TCCAATGAAG  
 801 TGGCCTGAGA ACCCCTCAGT AGACTCCACA GACAAGATAC  
 841 TGCCTGAAGT GGACCCCCAC CAGGICTCCC TCCCTGCCTC  
 881 CAGGICTCTA CAATCTATTG GCCACTGGGA ATGCTCAGGC  
 921 CAGATAAGCA CATTTCTTA GAGATGCGAC CCTGCCTTTC  
 961 TTGTCAATGC CTAGTAGTGG CTACCAGGAT CAATTGGTAA  
 1001 CAGGTCGTTT AATGTATAAT CAGAAGGACT TCCACTTCCT  
 1041 TCTTTCCTTT TATTTATTTA TTTATTTATT TACTTATTTA  
 1081 TTTATTTATT TTCTAGTTTC CTTTTCATGC TGGGACATCC  
 1121 CTTTGAGCAA AGGTGTGAAA GGGGAATAAT GGACCAGAAT  
 1161 TGCAAGGAAC ACAGGGAGGT CTTGAGGAAT GGGGGACACC  
 1201 AGAGCCAGTA GAAGGACATT CCTGGTGGCA AGGACCTGAG  
 1241 GACTCGAAAG TTGGAAAGGG GTTCAGAATG GGTGCTTTAG  
 1281 AATGGCAATT GGAATAAAAA TCTACCAAAG AAAGAACCGG  
 1321 GAATTGGACT AAATGGAATA AATGGGATAA ACTCATTTTG  
 1361 GTTTCATGTT AAGACTTTCT TTCTTACATC CGTCTCTTC  
 1401 TAATCTTGCA ATTTTATTTT TTTACACATA TTGTGGAGAT  
 1441 TTGGTGTGTT GGAGCATTTT TCTTTGAAA TAGATTTCAA  
 1481 TTTACACAAA CCGCTAAATT TCCTCATTAG CCACCGAAAA  
 1521 ACTCACACAA CACAAACACA CACACACACA CACACACACA  
 1561 CACACACACA CACACACAAA CACGGATACA CAAAGTGGAT



1601 TGTGTCTGAG AGCAAGCTGT AAGCATCTAG GCTTCCTCCT  
 1641 CAGTCTAGAA GTCTGATATC TGTACTGTGA AAGTGAGGCT  
 1681 TTTGTGTTTG GTTGTGATGT ATGCAGACAG GTTTCCTACA  
 1721 AACACTGGGG AGATCCCCTA GATCAGCTTC TGCGGGTCTT  
 1761 GGATGAAGGG TGTGGCCGAA CTGCAAGCGG CTCAGGCTGA  
 1801 TTCTCCCACC TTCGGGATAA TAAACAATCT GTAGCCCATC  
 1841 TCTGCCTGCC CCTTGTTCGC CACCTGTGCT GCAGTGATTC  
 1881 CTCACCTGCA CCCCAGATC ACTTCTTTGTC TCCCAGCAGG  
 1921 GATGCAGTTC TTTGACTGGG TTCTCCCATC CTGCCCCAAA  
 1961 GTGCCCTGGCT TTAGGAAGGG AAGCTGCCAA GCTAAGCTTC  
 2001 GGCTTCCACA TTCTCCATTC AGCGGTTGAG CTTGATAGGA  
 2041 GGTGGGCCCTT TCCCTCCCTCG AGTTCCCTAT TCTTCTTGGA  
 2081 GAATTGTTCGT GCGGTGGCAG GTAAGAAGT CTGCTACCTT  
 2121 CTGGTCCGTT TCACAGCCGG CCTACCCAAA GATTCAGCTC  
 2161 AAGGAGGCAA GGTCAGCGGT AGCTGTGGAG ATAAGGGCAG  
 2201 ATCTCTTGGG AGGGGCCCGC CTTGGCAGCA GTCCAAGAGA  
 2241 GCAGGCAGGG GAGTGGGCGG GCCCTCCACC CAGCGGTAGA  
 2281 TAAAAGGCCA GGGTGTAGCA CGCTGAGGCC CTGACAAGCA  
 2321 AAGGACTTGC TTGCTCTTTC CTCCTGGGGC GGAGGTCTGT  
 2361 TTTGGTAGTG GCAAGTGAAA GAAGTAGCTC CTGGCAAGTC  
 2401 TTCCCTTCGTT GCTTGGCAGG ATGGAAGGTC GCAGCCGTGT  
 2441 GGGAAAATC AGGGTGGCCC GGGAAATCCCG GCGCCCAAAG  
 2481 ACCCTTTTGA AAGGCATGCA GAAGGAGGCC CTTCTAAGTC  
 2521 GTTTCGCCAA GAGCCCTTCC CCAAGCTTCC GGGCCAGGGA  
 2561 GCAACTAGCC AGGGAAATCG GGCTGCCTGA GTCCCCAGTG  
 2601 CGGGTGTGGT TTCAGAACCG CCGAAGTCGT TCGGCGATGG  
 2641 TGAGGCGAGG CCCAGAATCC TCAGCGAGCC GAAGCGTCCT  
 2681 GGAGTCCCCG AAGCGGCCTG AAATGTCTGA TGCCAGGGCA  
 2721 CAAGGTAGAA GGAGCAGCCC GTCGGGACGT AGGAGGCCTC  
 2761 GCACTCGACT CAGCTCGACA CAGTTGGCCA TCCTGCTCCA  
 2801 GGCTTTTGAG AGGACGGTAC CCGGGCATTG CTGCCAGAGA  
 2841 AGAGCTC

## 2. Melléklet: A "b" típusú ismétlődő egység nukleotid szekvenciája:

1 GAGCTCGCCC AAAAGACCGA TTTACCGGAA GACACGATTC  
 41 ACATATGGTT TCAGAACAGA AGAGCACGCC ACCGTGCCGC  
 81 CTGCACGGGT AGAGTTCCGA GTCCAGATTC CCTTTCCTCG  
 121 CCTGAGCCCA ATGAGGCAGT GCACGGCAGC GAGACTGAGC  
 161 GTGCACAGGA CAGCTTGTTG CCACTGGCTG CTGCATCACC  
 201 TGTGGGCACA GGCACTTTGA CCATGACTTG CCCATCAACC  
 241 TCCTACCGAA TGTCCCTGGCC TGCAGAGACG GCACAGCCCT  
 281 CTCCACCAGA CGTCCCAGAG ACACACACCG TTGCATGGAA  
 321 CACCTGCGCT GGGCAGCTCT TTCTTGATGC AGTGATGAAA  
 361 GAGATCCAAG ATGACCAAGA TCTCCGACCC CTGCAGGACT  
 401 CTCTGGATCC CTACGATCAA TGGGAGCAGA TGCAGCCAAC  
 441 CCTTGAGCCC CCTCTGGACG AAGAAGTCTA TCAAGCTCTA  
 481 CTGGATGACC TCTGTGACAC ACACTGCAAG GCTTCATCGT  
 521 CTCATCACGG ATGCCGTCTT CTTCCAGAAA GCAAGACCCCT  
 561 TACACCAAAC CCCAACACAG GACAAAATGG ACCTCCAACG  
 601 CCCCTGCCCC TGTTCGGGAG CCCATGGACA GGAATCCCCT  
 641 GACCCTAAGC AACTGGGAAA AATCACCAGA AACAGGAAGT  
 681 CGGGTGTGGA AGGCCATCCC AGGAGGCCGA GGAAGTAGAAA  
 721 ATTCCCTGGC AGGAGGCGTT TGCTGGAGAC TAGAGGGAAA  
 761 TTTCCCCAGT CATAACACAGG AAGACATCTC TCCAATGAAG  
 801 TGGCCTGAGA ACCCCTCAGT AGACTCCACA GACAAGATAC  
 841 TGCCCTGAAGT GGACCCCCAC CAGGTCTCAC TCCCTGCCTC  
 881 CTGTCTCTAC ACTCTAATTGG CCACTGTGGA ACGCTTGGGC  
 921 CAGATAAGCA CATTTCCTG GAGATGCGAG TCTGCTTTTC  
 961 GTGTCAATGC CTAGTAGTGG CTACCAGGAT CCATAGGTAT  
 1001 CAGGTCATTT AATGTATAAT CAGAGGGACG TCCAATTTCT  
 1041 TCTTCTTTTC TATAGATTGA ATGATTGATT GATTGATTTT  
 1081 CTAGTTTCCT TTTTCATGCTG GGGCATCTGT TTGAGCAAAG  
 1121 GAGTGAAAGG GGAATAATGG ACCAGAATTG CAAGGAACAC  
 1161 AGGGAGGTCT TGAGGAATGG GGGACACCAC AGTCAGGAGA  
 1201 AGGAATTCCT TGGTGCCTAC GAGCTGAGGA CTGAAAAGAT  
 1241 GGATAGGGGT TCAGAAATGGT TGTTTTAGAA TGGCAATTCA  
 1281 AATAAAAATC TACAAAAGAA AGAACCGGGA ATTGGACTAA  
 1321 ATGGAATAAA TGGAATAAAT TCATTTTTGG TTCATGTTAA  
 1361 GACTTTCCTT CTACATCCG TTCTCTTTCT TCTAATCTTG  
 1401 CAATTTTTAT TTTTGTATTAT ATTGTGGAGA TTTGGGTTTT  
 1441 TGGAGCATTT CTCTTTAAA ATAGTTTTCA ATTTACATAA  
 1481 ACCGCTAAAT CCCCTCATTA GCCACCTACA AACACACACA  
 1521 ACACAAATAC ACACACACAC ACACACACAC ACACACACAC  
 1561 ACACACAAAC ACAGACACAC ACATGGGTAC ACCAAGGGGA



1601 TTGTGTCTGA GGGCAAGCTG TAAGCATCTA GGCTTCCTCC  
 1641 TCAGTTTACA AGTTTGATAT CTGTACTGTG AAAGTGAGGC  
 1681 TTTTCTGTTT GGTTCTGATG TATGCAGACA GGTTTACAAC  
 1721 AAACACTGGG GAGATCCCCT AGATCAGCTT CTGCAGGACT  
 1761 TAGATGAAGG GTGTGGCCGA ACTGCAAGCC GCTCAAGCTG  
 1801 ATTCTCCCAC TTCGGGATAA TAAACACTCT GTAGCCCATC  
 1841 TCIGCCIGCC CCTTGTTCG AACCTGTGCT GCCCTGATTC  
 1881 CTCACTTGCA CCTGTGATC ACTTATCACC CAACAGGGAT  
 1921 ATAGTTCTTC GACTGGGTTT TCCCATCGTG CCCAAAGGTG  
 1961 CCAGGCTTTA GGAAGGGAAG CTGCCAAGCT AAGCTTCGGC  
 2001 TTCCATATTC TCCATTACCC GGTGAGCTT GACATGAGGT  
 2041 GGGCCCTTCC TTCCTCCAGT TCCCTATTCT TCCTGGCGAA  
 2081 CTGTCTGGG GGTGGCAGTA AAGAAGACTG CTACCTTCTG  
 2121 ATCCGTTTCC CTGCCCGCCT ACGTGAAGGT TCGACTCAAG  
 2161 GAGGCAAGGT CAGTGGTAGC TGTGGAGATG AGGCCAAATC  
 2201 TCTTGGCAGG CGCCAGACGC GGCAGCAGTC CAAGAGAGCA  
 2241 AGCGGGGAGA GTGGGGCGGG CACCACCAG CGGTAGATAA  
 2281 AAGGCCAGGG TGTAGCACGC TGAGGCCAG ACAAGCAAAG  
 2321 GACTTGTCTG CTCCTTCTC CTGGGGCGGT GGCTGGTTT  
 2361 GGTAGTGGCA AGTGAAGGAA GTAGCTCTG GCAAGTCTTC  
 2401 CTTCGTTGCT TGGCAGGATG GAAGGTCCGA GCGTGTGGG  
 2441 AAAAATCAGG GTGGCCCGGG AATCCCGCG CCCAAAGACC  
 2481 CTTTTGAAAG GCATGCAGAA GGAGGCCCTT CTAAGTCGTT  
 2521 TCGCCAAGAG CCTTCCCCA AGCTTCCGGG CCAGGGAGCA  
 2561 ACTAGCCAGG GAAATCGGGC TGCCTGAGTC CCGAGTCCGG  
 2601 GTGTGGTTTC AGAACCGCCG AAGTCGTTCC GGGATGGTGA  
 2641 GCGAGGCCC AGAATCCTCA GCGAGCCGAA GCGTCTGGA  
 2681 GTCCCCGAG CCGCCTGAAA TGTCTGATGC CAGGGCACAA  
 2721 GGTAGAAGGA GCAGCCCATC GGGACGTGGA GGCTTCGCAC  
 2761 TCGACTCAGC TCGACACAGT TGGCCATCCT GCTCCAGGCT  
 2801 TTTGAGAGGA CCGTACCCGG GCATTGCTGC CAGAGAAGAG  
 2841 CTC



3. Melléklet: Az 1.2 kb méretű deléció ismétlődő egység nukleotid szekvenciája:

```

1      GGATCCACTT GACTGGAAAA TCTTTTCTCA ACATTTAGTT
41     CTTAGGCACC AACTGTCCTT GAAGATGAAG TGTATTTCTA
81     GTATGCAGCA GAAGGATGGA TTCTGTTTTT TCATTCATTC
121    CACTTACCTG TGTCTTTTTA TATGIGAGTT AAGTAGATTA
161    ATATTGAGAG ATACAAGAAC GATTGGITGA TCATTCCTCT
201    TTGTTTTTGA TGTTGGTGGT GTGGTGGTGG TGGTGGTGG
241    ATGGTGTGTG GATTTCCACC CAGTATTCCT TTGGCTGTGT
281    GTAAAGTGGG ATTATCTATT GCCTGTTTTT CTGGTTGTAG
321    TCAATTTCCCT TGGGTATGAG TTATACTTCC AACACTTTCT
361    GTATGGCTTG ATTGGTGGAT ATGTGTTGTA AATTTTTTGT
401    GAGITCAAGG TAGGGAACAA GATTGTTTTA CATGTAAGTC
441    CATTCCTCCCT CTCCCTCCCC CTAACTCCAT CCCACCTCGA
481    ATGATGCAGT GGACCACAGC TAGACTGAGA GTACACAGGA
521    CAGCTTGATG ACACTGGCTG CTGCAGTGTG CAAGAGAGCA
561    GGCAGGGGAG TGGGCGGGGC CTCCACCCAG CGGGTAGATA
601    AAAGGCCAGG GTGTAGCACG CTGAGGCCCA GACAAGCAAA
641    GGAATTGCTT GCTCTTTTCT CCTGGGGCGG AGGICTGTTT
681    TGGTAGTGGC AAGTGAAAGA AGTAGCTCCT GGCAAGTCTT
721    CCTTCGTTGC TTGGCAGGAT GGAAGGTCCG AGCCGTGTGG
761    GAAAAATCAG GGTGGCCCGG GAATCCCGGC GCCCAAAGAC
801    CCTTTTGAAG GGCATGCAGA AGGAGGCCCT TCTAAGTCGT
841    TTCGCCAAGA GCCCTTCCCC AAGCTTCCCG GCCAGGGAGC
881    AACTAGCCAG GGAAATCGGG CTGCCTGAGT CCCGAGTGGC
921    GGTGTGGTIT CAGAACCGCC GAAGTCGTTT GGGGATGGTGT
961    AGGCGAGGCC CAGAATCCTC AGCGAGCCGA AGCGTCTTGG
1001   AGTCCCCGAA GCCGCCTGAA ATGTCTGATG CCAGGGCACA
1041   AGGTAGAAGG AGCAGCCCGT CGGGACGTAG GAGGCCCTGC
1081   ACTCGACTCA GCTCGACACA GTTGGCCATC CTGCTCCAGG
1121   CTTTTGAGAG GACGTACCCG GGCATGTCTG CCAGAGAAGA
1161   GCTC

```

## ***KÖSZÖNET NYILVÁNÍTÁS:***

Köszönetemet nyilvánítom témavezetőmnek, Dr Hadlaczky Gyulának akinek támogatása lehetővé tette ennek a munkának a létrejöttét. Ugyancsak köszönettel tartozom Dr Cserpán Imrének, Dr Praznovszky Tündének és Kereső Juditnak akik az itt ismertett kísérletek egy részét végezték, illetve segítséget nyújtottak azok megvalósításában. Mindenek felett azonban köszönettel tartozom szüleimnek akiknek áldozatos munkája, szeretete és feltétlen támogatása nélkül ez a pillanat soha nem jöhetett volna el.