

## Bevezetés

### A $\lambda$ bakteriofág másodlagos *attachment* helyekre való integrációja megváltoztatott specifitású *int* mutánsok esetén

című doktori értekezés tézisei

Készítette: Szameczné Rutkai Edit

Témavezető: Dr. Dorgai László

Bay Zoltán Alkalmazott Kutatási Közalapítvány

Biotechnológiai Intézete

Szeged, 2006

Napjainkban az egyik legjobban ismert vírus - baktérium rendszer az *Escherichia coli* és vírusa, a  $\lambda$  bakteriofág. A  $\lambda$  más temperált bakteriofágokhoz hasonlóan kétféle stratégiát követhet, miután bejutott a gazdaszervezetbe. A lítikus út során a fág lízissel elpusztítja a baktériumsejtet, miközben annak erőforrásait kihasználva, megközelítőleg száz új vírusrészecske képződik. A másik lehetőség a lizogén út, ekkor a fág DNS beépül a gazda kromozómájába.

A beépülés (inzerció) a helyspecifikus rekombináció révén valósul meg, melynek apparátusa egy-egy egyedi szekvenciát felismerve a baktérium kromozómán és a fág DNS-en, egy reciprok rekombinációval fúzionálja a két genomot. A beépült fág genomot hívjuk profágnak.

A rekombináció kulcsenzime az Integráz, amely katalizálja a DNS szálak hasítását, a szálkicserélődést, majd a kicserélődött szálak újraligálását.

Azokat a szekvenciákat, melyeket a  $\lambda$  Integráz felismer és amelyeknél megtörténik a rekombináció, nevezzük *attachment (att)* helyeknek.

Az elsődleges *att* helyeken kívül mind a baktérium, mind a fág tartalmaz ún. másodlagos *att* helyeket. A beépülés hatékonysága ezeken a helyeken nagyságrendekkel kisebb, mint az elsődleges *att*-helynél. Sosem tapasztaltak másodlagos helyre történő integrációt, ha az elsődleges *attB* hely is megtalálható a genomon. A másodlagos helyek felépítése megfelel az elsődleges helyek felépítésének, szekvenciájuk több-kevesebb hasonlóságot mutat az elsődleges *att* szekvenciához.

A HK022 fág a  $\lambda$ -hoz hasonlóan temperált bakteriofág. A két rendszer eddigi vizsgálatának minden eredménye arra utal, hogy a két fág azonos reakcióutat követ a beépülés és a kivágódás során. Viszont a strukturális, szabályozási és reakciómechanizmusbeli hasonlóságok ellenére egyik sem katalizálja a másik fág beépülését illetve kivágódását kellő hatékonysággal.

A HK022 Int és a  $\lambda$  Int közötti hasonlóság alapján tanulmányozhatók az evolúciós változások mechanizmusai a fehérje-nukleinsav kölcsönhatásokban. A két fehérje az evolúció során eljutott egy olyan pontig, ahol a rekombinációs helypreferenciájuk különböző és könnyen mérhető, de még nagyfokú hasonlóságot mutatnak szekvenciáikban, és alacsony szintű keresztreakció is megfigyelhető.

A kérdés az, hogy hogyan alakulhat ki egy új specifitás egy már meglévőből úgy, hogy a folyamat során a rendszer még időlegesen se váljon funkcióképtelenné. Az Int és felismerőhelye közti kölcsönös funkcionális kompatibilitás elvesztése ugyanis csökkentené a lizogénizáció hatékonyságát.

A dolgozat kiindulópontjaként szolgáló korábbi munka során ismertté váltak a  $\lambda$ /HK022 fágok helyspecifikus rekombinációs rendszereiben az eltérő specifitásért felelős elemek. Azonosításra kerültek az Integrázok specifitás-különbségéért felelős aminosavak, és a *core* típusú Int-kötő szekvenciákban azok a nukleotidok, melyek a nem saját Int által katalizált rekombinációt gátolják. A vad típusú és mutáns Integrázok valamint a vad típusú és mutáns *att* szekvenciák különböző kombinációkban, *in vivo* mért rekombinációs tulajdonságai alapján valószínűsíteni lehetett egy utat, melyet követve egy létező *att*/Int rendszer specifitása megváltozhat az evolúció során. Eszerint egy vagy több Int mutáció relaxálja a rekombináz specifitását, majd további mutációk beszűkítik azt úgy, hogy az eredeti szekvenciát már nem rekombinálja hatékonyan, de egy másikat már igen. Később újabb mutációk, mind az Int-ben, mind az új *core* szekvenciákban finomhangolják a rendszert.

## Célkitűzés

A hipotézis megfogalmazza a specifitás változásának egy lehetséges mechanizmusát, de arra nézve nem ad választ, hogy ez a valóságban hogyan következhet be. Az ugyanis a rendszer komplexitása és a szükséges mutációk relatíve nagy száma miatt valószínűtlen, hogy a kísérletesen megvalósított  $\lambda \rightarrow$  HK022 irányú specifitásváltozás reális körülmények között is hasonló módon bekövetkezzen anélkül, hogy a rendszer funkcióképessége a teljes folyamat során megmaradjon.

A relaxált specifitású Int-nek találnia kell egy vagy több olyan szekvenciát, melyet relatíve jó hatékonysággal rekombinálni képes. Az ilyen szekvenciák jó jelöltjei a gazdabaktérium kromoszómáján található másodlagos *att* helyek, melyek szekvenciái több-kevesebb mértékben hasonlítanak az *attB*-re, és a vad típusú Int kis hatékonysággal ugyan, de képes rekombinálni őket. Feladatomban annak vizsgálata volt, hogy a másodlagos *attB*-k szerepet játszhatnak-e a specifitás változásában, és ha igen, akkor milyen mechanizmus szerint.

## Eredmények és diszkusszió

Első lépésként meghatároztuk a vad és a mutáns Integrázt hordozó fágok integrációs képességét az *Escherichia coli*-ban található másodlagos helyek összességénél.

*Southern* hibridizációs kísérletben kimutattuk az egyes fágok által használt másodlagos helyeket.

Az egyoldalú PCR módszerével azonosítottuk ezen másodlagos helyek nagy részét. Összesen 19, eddig nem leírt másodlagos *attachment* hely szekvenciáját határoztuk meg. A szekvencia ismeretében az általunk izolált másodlagos helyek nagyrésze beazonosítható volt a lizogén populációk *Southern* hibridizációs vizsgálatánál kapott képen.

A következőkben 459 egyedi lizogént vizsgáltunk meg, hogy esetükben melyik másodlagos helyen történt meg az integráció. Így kvantitatív képet kaptunk az egyes helyek használatáról.

A másodlagos helyek szekvenciaanalízisének legfeltűnőbb eredménye az overlap régió első nukleotidjainak erős konzerváltsága. Az eddig elfogadott nézet szerint, az overlap régiók konkrét szekvenciája indifferens, de a rekombinálódó partnereket tekintve azonosnak kell

lennie. A kísérleteinkben viszont ez az azonosság csak az overlap bal oldali régiójánál volt megfigyelhető.

197 egyedi lizogén vizsgálatával bebizonyítottuk, hogy ha az integrációban résztvevő két *overlap* régió jobb oldali szegmense több nukleotidban is különbözik, akkor az integráció az irányított szegregáció szabálya szerint történik, azaz az *attL* régió overlap-je az *attP*-től, míg az *attR* régió overlap-je mindig az *attB*-től öröklí a szekvenciáját. Erre magyarázat lehet, hogy az overlap régiók jobb oldalának szekvencia különbsége miatt a Holliday struktúra nem vándorol el az overlap régió végéig, a második szálcsere nem a 4-5 nukleotidok között történik, hanem megoldódik hamarabb, feltételezhetően a baktérium homológ rekombinációs rendszere révén.

Eredményeink alapján felállítottuk a specifitászváltozás *chromosome jumping* modelljét. A modell feltételezi egy *attR* transzdukáló fág képződését, amelyben az *attR* döntő mértékben megőrzi az *attP* strukturális komplexitását, ami biztosíték arra, hogy néhány mutáció után az integratív funkció gyorsan helyreállhasson.

A modell két elemét is teszteltük, egyik kísérletben mesterséges plazmidszubsztrátok, míg a másikon egy *attR* transzdukáló fág vizsgálatával. A transzdukáló fág esetében bizonyítottuk, hogy

struktúrája megfelel a modell által előrejelzettnek, integrációs és kivágódási funkciói épek. A plazmidszubsztrátok esetében mértük az integrációs és a kivágódási hatékonyságot különböző *att* szekvenciák esetén. Érdekes és nem várt megfigyelés a bizonyos szubsztrát/Int kombinációkban észlelhető szignifikáns különbség az integratív és kivágódási reakciók hatékonyságában. Mindkét kísérletben kapott eredményeink alátámasztják a modellünket.

Célunknak megfelelően, sikerült megvizsgálunk a másodlagos helyek specifitászváltozásban betöltött szerepét. Felállítottunk egy modellt, amelynek több elemét kísérletesen is bizonyítottuk. Természetesen nem állítjuk, hogy a vizsgált fehérjemutációk és másodlagos helyek valaha is szerepet játszottak egy konkrét specifitászváltozásban, de eredményeink meggyőzően bizonyítják a folyamat lehetséges voltát.

## Alkalmazott módszerek

A standard mikrobiológiai és molekuláris biológiai módszereken túl, a következő módszereket használtuk kísérleteink során:

- a  $\lambda$  fág *E. coli* sejtekbe való integrációjának hatékonyságát *in vivo* mértük, a fágfertőzést követően izolált lizogén baktériumtelepek számát viszonyítottuk az összes túlélő telep számához,
- plazmidszubsztrátokon az integráció és a kivágódás frekvenciáját a *lacZ* gén expressziójának változásával, X-gal szubsztrát jelenlétében a színreakciót adó telepek arányával határoztuk meg,
- az egyoldali PCR módszerével azonosítottuk a leggyakrabban használt másodlagos helyeket, majd ugyanezen módszer segítségével ezen másodlagos helyek használatának gyakoriságát is meghatároztuk.

## Summary

During the project described in the theses, I examined the possibility of the change of the specificity in the site-specific recombinational systems of the lambdoid temperate bacteriophages. One of the most studied site-specific recombinational system is that of the *Escherichia coli* and the phage  $\lambda$ , that following recognition of the unique sequences on the bacterial chromosome and phage DNA fuses the two genomes by a reciprocal recombination.

This fusion is catalyzed by Integrase (Int), a phage encoded site-specific recombinase. The sequences where the recombination takes place and are identified by the Int are called *attachment (att)* sites. Beyond the primary *att* sites both the bacterial chromosome and the phage contain so-called secondary *att* sites. The recombinational efficiency of these sites is lower by several orders of magnitude than that of the primary *att* sites.

The phage HK022, similarly to the phage  $\lambda$ , can follow the lysogenic way after infecting the bacterial cell. Former studies comparing the site-specific recombinational systems of two phages revealed that the mechanism of both the integration and the excision is

very similar. In spite of the structural and regulational analogy, none of the two Integrases can catalyze efficiently the integration or excision of the other phage.

Studying the differences and similarities of the HK022 Int and the  $\lambda$  Int, the mechanisms of the evolutionary changes in the protein-DNA interactions can be understood. During former studies leading to my project, the elements responsible for the different specificity in the two site-specific recombinational systems of these phages was identified. Based on the *in vivo* recombinational efficiency of different combinations of the wild type or mutant Integrases and the wild type or mutant *att* sites we hypothesized a pathway that describes a possible way how the specificity of an existing *att/int* system can change during the evolution: one or more Int mutations relax the specificity of the recombinase and further mutations restrict it, resulting in that the original sequences will not be recombined efficiently but a new one will be. Finally, new mutations in the Int and in the new Int-binding sites will refine the system.

My specific aims were to investigate if the secondary *att* sites could play a role in the change of the specificity of a recombinational system and to determine the mechanism of this possible pathway.

The efficiency of the integration into the *Escherichia coli* chromosome using all of the secondary *att* sites *en masse* was measured using  $\lambda$  phages carrying the wild type or mutant Integrases, and the ratio of the integrational efficiencies at primary and secondary *attB* sites was determined.

In Southern hybridization assay, the secondary sites used most frequently by phages carrying the wild type and mutant Int was detected. The *att* regions of prophages integrated into secondary sites were amplified in one-sided PCR using the population of thousands of single lysogen strains as a template. By determining the sequence of these PCR products, 19 new secondary *attachment* sites were identified.

The relative frequency of the utilization of these secondary sites was estimated by the amplification of the prophage *att* sites using genomic DNA isolated from large lysogen populations (~3000 individual colonies) as a template. The recombinational efficiency at the most frequently used secondary sites was measured similarly, using 459 individual lysogen.

The most outstanding result of analysis of the secondary *att* sites is the strong conservation of the first three nucleotides of the *overlap* region. Previous data showed that the sequences of the *overlap* regions

in the recombining *att* sites should be identical, but the sequence itself is indifferent. Our experimental data proves that this identity is important only for the first three nucleotides of the *overlap* regions in the case of integrational events at the secondary sites.

By examining 197 independent insertions we demonstrated that unidirectional segregation is the rule when the right sides of the two *overlap* regions are multiply mismatched, that is that all of the *attL* overlap regions had the sequence of the  $\lambda$  and all of the *attR* overlap regions had the sequence of the secondary sites. This observation can be explained by the inhibition of the second strand swap because of the differences in the sequences of the right side of the *overlap* regions, so the Holliday structure created by the first strand exchange will be resolved not by the Int but some other mechanism for example by the bacterial homologous recombination or replication.

The results presented in this work led us to set up a model called chromosomal jumping. We have examined two predictions deduced from our model. In the first experiment, the recombinational efficiencies were measured by artificial plasmid substrates containing secondary *att* sites. In the second, the integrational ability of an *attR* transducing phage

was tested. The results from these experiments are compatible with the mechanism described in the model and confirm the availability of it.

The possible role of the secondary *att* sites in the change of specificity of the site-specific recombinational system was investigated according to our initial aims. A model was set up and some elements of it were proved by experimental data. Of course, we cannot say that the examined mutations of the Int protein and the secondary sites were playing role in the change of specificity, but our results confirm its possibility.

#### **A fokozat megszerzéséhez felhasznált tudományos közlemények**

- E. Rutkai**, L. Dorgai, R. Sirot, E. Yagil and R. A. Weisberg: Characterization of secondary attachment sites used by phage  $\lambda$ , and a role for such sites in changing phage insertion specificity  
*J.Mol.Biol.* 329 (5) 983 – 996, 2003
- E. Rutkai**, A. György, L. Dorgai and R. A. Weisberg: The Role of Secondary Attachment Sites in Changing the Specificity of Site-Specific Recombination  
*J. Bacteriol.* 2006 (közlésre elfogadva)

## Egyéb tudományos közlemények

T. Bubán, **E. Rutkai**, L. Dorgai and Thomson, S. V.: Prediction of infection risk on the basis of weather-related factors and *Erwinia amylovora* colonization in apple and pear flowers  
*Int. J. Hort. Sci.* 10 (2), 39-54, 2004

S. Sloan, **E. Rutkai**, R. King and R.A. Weisberg: Persistence of RNA mediated antitermination (kézirat)

## Előadások

**E. Rutkai**, L. Dorgai, R. Sirot, E. Yagil and R. A. Weisberg: Conserved elements in secondary attachment sites used by  $\lambda$  and *int* mutants with altered insertion specificity.  
*Analytical Genetics*, Santorini, Greece, Oct. 1-6, 2002.

**E. Rutkai**, S. Sloan, R. King and R. A. Weisberg: Persistence of RNA mediated antitermination.  
*The New Phage Biology*, ASM Conferences, Key Biscayne, Florida, USA, Aug. 1-5, 2004

## Poszterek

**E. Rutkai**, L. Dorgai, R. Sirot, E. Yagil, and R.A. Weisberg: Secondary attachment sites for phage  $\lambda$  and the evolution of phage insertion specificity  
*Molecular Genetics of Bacteria and Phages*, Cold Spring Harbor, USA, Aug. 20-25, 2002.

## Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani

- témavezetőmnek, Dr. Dorgai Lászlónak, hogy hasznos tanácsaival irányította kutatásaimat és elősegítette szakmai fejlődésemet, karrierem építését,
- munkatársamnak, György Andreának, aki eredményes munkájával segítette doktori dolgozatom elkészítését,
- Bálint Arankának és Vörös Andreának, akiknek a mindennapi segítségére és értékes munkájára mindig számíthattam,
- Dr. Kiss Istvánnak és Urbán Gabriellának, akik hozzáértő kritikájukkal sokat segítettek dolgozatom megírásában,
- Dr. Kálmán Miklósnak, a Bay Zoltán Alkalmazott Kutatási Közalapítvány Biotechnológiai Intézetének igazgatójának, hogy a kutatásaimhoz szükséges feltételeket biztosította,
- Dr. Robert Weisbergnek akitől sokat tanulhattam és aki nagyban hozzájárult dolgozatom elkészültéhez.