Maria Elisa Manetti

Caracterização de retrotransposons similares a Tnt1 em espécies silvestres e cultivadas do gênero *Solanum*

São Paulo

2007

Maria Elisa Manetti

Caracterização de retrotransposons similares a Tnt1 em espécies silvestres e cultivadas do gênero *Solanum*

Tese apresentada ao Instituto de Biociências da Universidade de São Paulo, para a obtenção de Título de Doutor em Ciências, na Área de Botânica.

Orientadora: Prof. Dra. Marie-Anne Van Sluys

São Paulo
2007

I - INTRODUÇÃO

I.1 – Descoberta dos elementos de transposição

Bárbara McClintock descreveu a existência dos elementos de transposição (TEs) na década de 40, estudando grãos de milho variegados. Identificou dois loci dominantes que interagiam entre si denominados *Dissociator* (Ds) e *Activator* (Ac). Em 1948, fez a descoberta surpreendente de que Ds e Ac podiam se transpor, ou seja, mudar de posição no cromossomo. Os efeitos da transposição foram constatados através da mudança na coloração do pericarpo dos grãos de milho, e a quebra do braço curto do cromossomo 9. Analisando cruzamentos entre linhagems controladas, McClintock concluiu que Ac controla a transposição de Ds e que o movimento do último promovia a ruptura do cromossomo. Quando Ds se excisava daquela posição, o gene responsável pela coloração na camada de aleurona era liberado do efeito supressor, e voltava ao estado nativo permitindo a síntese do pigmento. O efeito de mosaico de alguns grãos deve-se à coexistência de células contendo o gene da coloração inativado pela inserção de Ds e células com o gene ativo (figura 1). Entre 1948 e 1950, McClintock desenvolveu uma teoria postulando que os elementos móveis (transponíveis) regulam os genes, inibindo ou modulando sua ação. Denominou Ds e Ac como “elementos controladores” para distingui-los dos genes. A descoberta de McClintock desafiou o conceito de que o genoma é um conjunto estático de informações transmitidas entre gerações. Em decorrência destas descobertas, Bárbara McClintock foi agraciada com o prêmio Nobel de Fisiologia e Medicina em 1983.

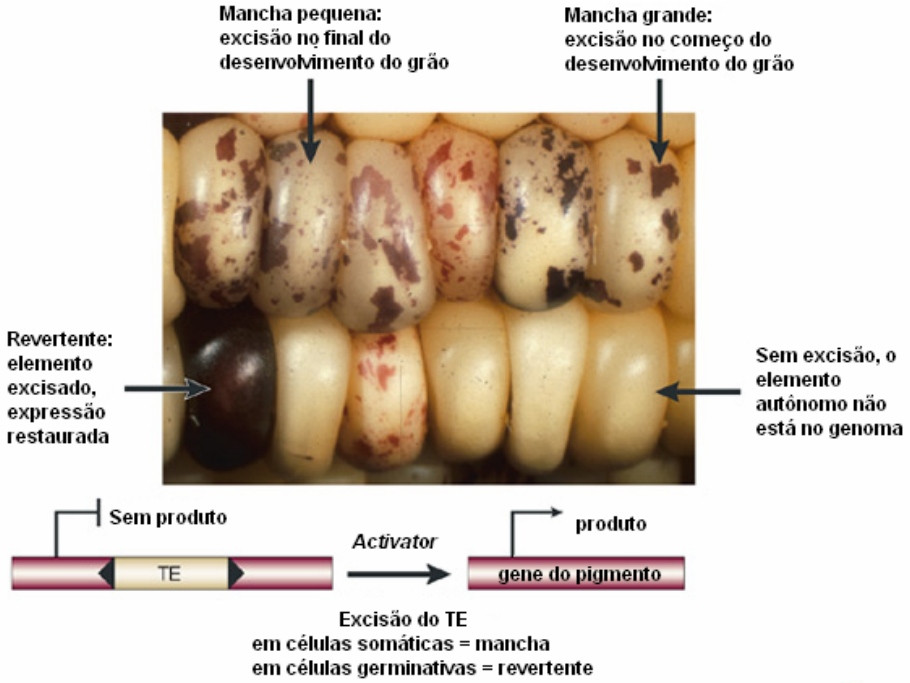

Figura 1. Diferentes fenótipos observados em grãos de milho devido à ação de TEs.
A instabilidade dos fenótipos se deve à inserção ou excisão do elemento em um gene que codifica uma enzima que participa da rota de biossíntese do pigmento antocianina. Setores revertentes (com pigmento) resultam da excisão do elemento em células da camada de aleurona. O tamanho do setor reflete o momento durante o desenvolvimento onde a excisão aconteceu. Adaptado de Feschotte *et al.* 2002.

I.2- Elementos de transposição

Atualmente, caracterizam-se os TEs como segmentos de DNA que possuem a capacidade de catalisar seu próprio movimento e inserir-se em novas regiões do genoma. Como conseqüência da sua atividade os TEs podem gerar alterações funcionais nos genes aos quais se associam. Quando inseridos dentro da região codificante de um determinado gene podem inativá-lo ou alterar seu produto. Quando inseridos dentro ou próximo a regiões promotoras podem alterar os padrões de expressão do gene (Kashkush *et al.* 2003). A atividade dos TEs pode ainda ser responsável pela mobilização de outras porções do DNA durante a transposição, ou de mecanismos de recombinação ilegítima. Já os

elementos replicativos apresentam como conseqüência do seu modo de transposição o aumento do tamanho do genoma hospedeiro (Kidwell, 1997; Fedoroff, 2000). O seqüenciamento de vários genomas evidenciou que os TEs são o componente majoritário dos genomas de eucariotos, compondo pelo menos 45% do genoma humano (Lander *et al*. 2001) e de 50 a 80% dos genomas em algumas gramíneas (San Miguel *et al*. 1998, Vicient *et al*. 1999).

Os TEs podem ser classificados em duas classes dependendo do intermediário de transposição, que pode ser uma molécula de RNA (classe I) ou de DNA (classe II) (Voytas *et al*. 1992, Feschotte *et al*. 2002). Os elementos da Classe I ou retroelementos possuem um mecanismo replicativo de transposição, que produz como resultado de sua mobilização um aumento do número de cópias e, concomitantemente, um aumento no tamanho do genoma hospedeiro (San Miguel and Bennetzen, 1998).

Os retroelementos podem ou não possuir longas terminações repetidas (LTRs). Nos retrotransposons com LTRs as mesmas localizam-se em suas extremidades na orientação direta. A LTR é subdividida nas regiões U3 (3' RNA), R (RNA repetido) e U5. (5' RNA) (figura 2). Os sinais regulatórios tais como o região promotora de transcrição, regiões de término de transcrição e os locais da poliadenilação estão localizados nas LTRs (Kumar and Bennetzen 1999). Além da importância funcional destas seqüências, diversos estudos relataram que as LTRs são as regiões que evoluem mais rapidamente nos retrotransposons (Kalmykova and Gvozdev 2004; Vicient *et al.* 2005). Os elementos com LTRs codificam uma proteína (*gag*) similar a um capsídeo viral e uma poliproteína (*pol*). Esta, por sua vez, quando processada resulta em várias proteínas maduras: a protease (PROT), que está envolvida no processo de maturação da poliproteína; a integrase (INT), necessária para a integração da cópia de DNA no genoma hospedeiro; e a proteína transcriptase

reversa/RNAseH (RT/RNAseH), responsável pela síntese do cDNA a partir do mRNA (figura 2). De acordo com a posição da integrase na poliproteína e o grau de semelhança entre as seqüências, os retrotransposons são divididos em duas famílias: Ty1-cópia e Ty3-gypsy (Boeke and Corces, 1989; Zachar *et al*. 1985) (figura 2).

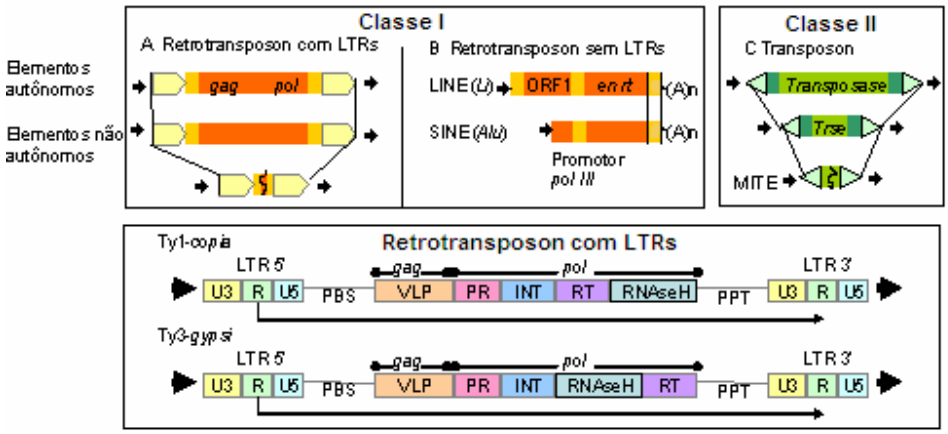

Figura 2. Estrutura geral dos Elementos de Transposição.

A Classe I ou retroelementos é composta por: retrotransposons com LTRs; sem LTRs; e retrovírus (não representado). Nos retrotransposons com LTRs (B), a orientação das mesmas é de maneira direta (triangulo amarelo claro). Elementos autônomos possuem dois genes: *gag* e *pol* (caixas laranja). A maioria dos elementos não autônomos não possui toda a região codificante. Os retrotransposons sem LTRs (B) são divididos em LINEs e SINEs. A região codificante inclui: ORF1 (similar a *gag*); *en* (endonuclease); e *rt* (reverse transcriptase) genes (caixas laranjas). A Classe II está composta por transposons de DNA com TIRs (terminal inverted repeats) (triângulo verde claro) e um gene (*transposase*) que codifica para uma transposase (caixa verde). Elementos não-autonomos são geralmente derivados de elementos autônomos por uma deleção. Caixa inferior: estrutura mais detalhada de um retrotransposon com LTR. Cada LTRs é dividida em três regiões: U3, R, e U5, que apresentam as sinais para a iniciação e terminação da transcrição. O transcrito (seta fina) é mostrado abaixo de cada bloco, começa no final 5' da região R da LTR'5 e termina na posição '3 da região R pertencente a LTR'3. Os genes do retrotransposon codificam varias proteínas: proteína do capsídeo viral (VLP), integrase (INT), protease (PR), transcriptase reversa (RT) e ribonuclease H (RNAase-H). Outras regiões características das seqüências são: PBS (*primer binding sites*), PPT (*polypurine tracts*). Pol III B sitio de reconhecimento da RNA polymerase III. Setas: direção do elemento. Adaptado de Kumar *et al*.1999 e Feschotte *et al.* 2002.

Os retrovírus têm estrutura similar a um retrotransposon com LTRs, com a presença adicional de uma ORF contendo o domínio *env*, que lhes confere o caráter infeccioso (Xiong and Eickbush, 1990). Já os retroelementos sem LTRs compreendem os LINEs (long interspersed nuclear elements) e SINEs (short interspersed nuclear elements). Os LINEs codificam duas ORFs, que são transcritas como um mRNA bicistrônico composto de uma ORF1 (proteína que se une ao RNA) e uma ORF2 (endonuclease e atividade de transcriptase reversa). Os LINEs e os SINEs terminam em uma seqüência simples repetitiva, geralmente uma cauda poli-A (figura 2). Os transcritos dos LINEs iniciam-se a partir de um promotor dentro da extremidade 5' do elemento e terminam freqüentemente junto à seqüência simples repetitiva. Os SINEs são caracterizados por um promotor interno de pol III de RNA próximo a extremidade 5′. Os SINEs são um grupo heterogêneo de elementos, que variam no comprimento de 90 a 300 pb e são derivados de uma variedade de genes de tRNA ou do RNA 7SL. Os SINEs são considerados elementos defectivos que parasitam a maquinaria de transposição dos LINEs (Ogiwara *et al*. 1999; Wessler 2006).

Muito pouco se conhece sobre a origem dos elementos da Classe I nas plantas. A ubiqüidade dos retrotransposons nas plantas e a sua heterogeneidade indicam que estiveram presentes nos primeiros vegetais, mas evoluíram variando seu número de cópias, localização, e estrutura das seqüências (Kumar and Bennetzen, 1999). Os TEs de Classe II, ou transposons clássicos, codificam geralmente para uma única proteína, a transposase, responsável pela excisão e integração do elemento. Possuem em suas extremidades repetições invertidas terminais (TIRs)(Feschotte *et al*. 2002) (figura 2).

Cada família de TEs contém elementos autônomos e não autônomos. Os elementos autônomos têm ORFs que codificam os produtos requeridos para a transposição. Já os

elementos não autônomos não codificam proteínas da transposição, mas podem se transpor porque retêm as seqüências *cis* necessárias para a transposição. A integração da maioria dos TEs resulta na duplicação de uma seqüência genômica curta (TSD: do inglês, *target site duplication*) no sitio de inserção (Wessler 2006) (figura 2).

I.3 – Ciclo de vida de um retrotransposon

O ciclo de vida do retrotransposon resulta na amplificação do genoma e por esta razão sua atividade é controlada finamente pelo hospedeiro. Além dos ciclos de amplificação, o tamanho do genoma do hospedeiro pode diminuir devido a eventos de recombinação homóloga dentro de um elemento ou entre elementos homólogos.

O processo de replicação de um retrotransposon é chamado retrotransposição porque o RNA é reverso-transcrito em DNA (oposto ao especificado pelo Dogma central da Biologia). Acredita-se que o ciclo de vida de um retrotransposon compartilha muitas características com o ciclo de um retrovírus tal como o da AIDS (Lentivirus). Ambos os ciclos de vida são mutagênicos e propensos a erros, já que a transcriptase reversa tem baixa atividade de processamento quanto à incorporação de bases errôneas não complementares à fita molde de RNA. Há também a possibilidade de recombinação durante a transcrição reversa. Uma vez que uma cópia replicada é introduzida no genoma as mutações neutras no decorrer do tempo irão gerar polimorfismos nas seqüências das LTRs, sinais de processamento e regiões codificantes, incluindo o surgimento de novos codons de parada de tradução (*stop codons*). Portanto, os retrovírus e os retrotransposons geram populações de seqüências similares denominadas *quasispecies* (Casacuberta *et al.* 1995).

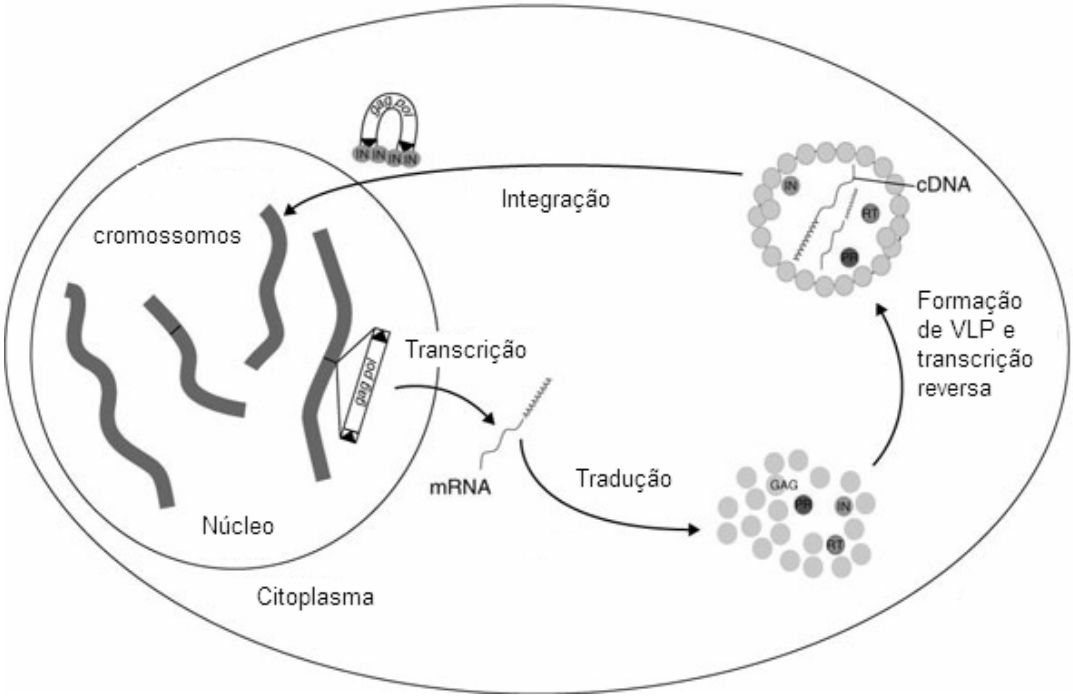

Figura 3. Ciclo de vida de um retrotransposon com LTRs.

IN: integrase; PR: protease; RT: transcriptase reversa; VLP: partícula tipo viral (do inglês *virus-like particle*). Triângulos pretos representam LTRs. Adaptado de Havecker *et al*. 2004.

O ciclo de vida de um elemento envolve primeiramente a transcrição do mesmo pela RNA polimerase II celular que reconhece o promotor localizado na região 5'LTR. O RNA é traduzido no citoplasma para formar as proteínas que formam as VLP (*virus-like particles*) e as que levam a cabo a transcrição reversa e integração da cópia de cDNA (Havecker *et al*. 2004) (figura 3). A transcrição reversa é um processo complexo e com muitas etapas que acontece dentro das VLPs no citoplasma (Kazazian 2004) (figura 4).

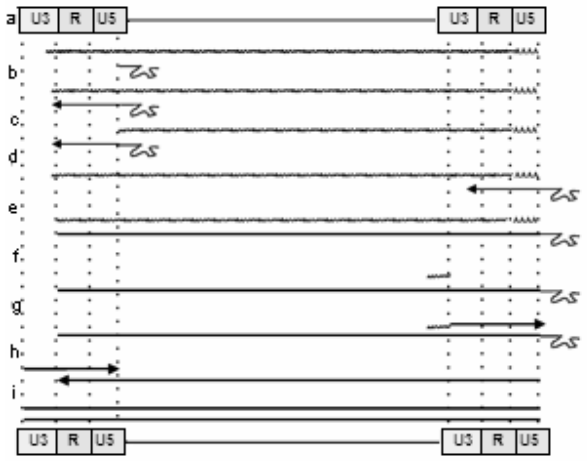

Figura 4. Transcrição reversa de transposons com LTRs.

A transcrição reversa começa com a cópia do RNA a DNA próximo do final 5' do RNA usando um tRNA como iniciador (a e b), seguido da degradação do RNA 5'(c). O novo DNA muda para a posição final 3' do RNA (d) e se completa a síntese da primeira fita de DNA (e). A RNAseH codificada pelo elemento degrada a maior parte do RNA (f). Uma pequena porção do RNA é utilizada como iniciador para a síntese da porção a direita do elemento (g). Uma segunda mudança de posição da segunda fita de DNA permite completar a síntese da segunda fita de DNA (i). Durante o processo as LTRs são formadas. Caixa colorida: LTRs e linhas verticais pontilhadas indicam regiões das mesmas (U3, R e U5); línea seguida de A: mRNA; línea irregular: tRNA; seta preta: direção da síntese. Adaptado de Kazazian 2004.

I.4 – O impacto dos TEs na evolução dos genomas

Por que os organismos mantêm uma alta proporção de TEs nos seus genomas? Eles cumprem alguma função? Como teriam influenciado a evolução da vida?

Vários processos participam da evolução dos genomas eucarióticos, tais como quebras cromossômicas e sua posterior translocação a outro cromossomo, duplicações de genes e segmentos, reordenação de domínios funcionais tais como os exons, criação de novos genes e conversão gênica. Durante as últimas décadas muitos destes processos foram

relatados como estando associados a TEs (Babushok *et al*. 2007; Cordaux *et al.* 2006; Deininger *et al*. 2003).

Desde o descobrimento dos TEs por Bárbara McClintock, muito tem se especulado e questionado sobre os TEs e seu possível papel na evolução dos genomas. McClintock chamou os TEs de "elementos controladores" ao observar mudanças nos fenótipos de grãos de milho devido à mobilização de TEs, propondo que controlariam o desenvolvimento normal do milho. Esta idéia foi rejeitada e anos depois ela propôs que os TEs seriam parte de uma resposta global do organismo ao estresse. Com a caracterização do primeiro TE em bactéria (Jordan *et al*. 1967) e o posterior reconhecimento dos mesmos como sendo componentes antigos dos genomas eucarióticos e procarióticos, especulações surgiram em relação ao papel dos TEs na evolução dos genomas.

Duas publicações (Doolittle and Sapienza 1980; Orgel and Criel 1980) postularam que a seleção natural, que opera dentro dos genomas, resultará inevitavelmente no surgimento de fragmentos de DNA sem nenhuma expressão fenotípica cuja única função é sua sobrevivência dentro dos genomas. Os TEs podem ser vistos como tais DNAs sem função fenotípica ou evolutiva. A partir destas informações surgiu a teoria dos TEs como parasitas (*selfish* DNA) ou DNA lixo (*junk* DNA) e a aceitação de que os TEs são componentes dos genomas sem um papel importante na evolução.

A teoria do DNA parasita foi amplamente aceita até a constatação de que os TEs geralmente perfazem a maior fração dos genomas dos organismos eucariontes. Estas revelações conduziram a novas questões: como os organismos e seus TEs coexistem? (Bowen and Jordan 2002).

O conteúdo de TEs varia de espécie para espécie tanto em relação às classes de TEs presente quanto quanto em relação à proporção no genoma. A levedura (*Saccharomyces*

cerevisiae) tem somente retrotransposon com LTRs (chamados elementos *Ty*), já nos genomas dos mamíferos predominam retrotransposons sem LTRs e os transposons da Classe II compõem menos de 5% da fração total de TEs. Os genomas das plantas com flores possuem uma grande variedade de ambas as classes de elementos, sendo os retrotransposons com LTRs a maior fração dos genomas caracterizados até agora. O potencial de amplificação dos retrotransposons com e sem LTRs foi primeiramente ilustrado em pesquisas com o gênero *Lilium.* As 14 espécies analisadas possuem genomas que variam entre 30.000-45.000 MB, cujo tamanho pode ser explicado por amplificações massivas de retrotransposons (Leeton *et al*. 1993). Diversos estudos mostraram que a amplificação de retrotransposon com LTRs esclarece a maior parte do paradoxo do número C entre importantes espécies de gramíneas. O paradoxo do valor C é a falta de correlação entre o tamanho do genoma e a complexidade alcançada pelo organismo. Na família das gramíneas a fração do genoma compreendida por retrotransposons com LTRs parece estar correlacionada com o tamanho de genoma: no arroz que apresenta o menor genoma (430 Mb), os retrotransposons com LTRs somam 15% do genoma (Jiang *et al.* 2001), por sua vez no genoma do milho (2.800 Mb) os retrotransposons representam 50-80 % (San Miguel and Bennetzen, 1998), e na cevada (4.800 Mb) estes constituem mais de 70 % do genoma (Vicient *et al*. 1999). A variação do conteúdo de TEs entre espécies pode ter influenciado tanto a qualidade quanto a quantidade da variação genômica, e como conseqüência a evolução de seus genomas. Dois estudos muito interessantes permitiram aprofundar o conhecimento na dinâmica dos genomas, particularmente nas gramíneas. Bennetzen e colaboradores (Science,1996) analisaram uma região de 280 kb em torno do gene de milho *adh1* e encontraram retrotransposons com LTRs uns dentro dos outros (*nested*). Posteriormente, conseguiram demonstrar que expansões dos retrotransposons com LTRs

dobraram o tamanho do genoma de milho dentro dos 6 milhões de anos depois da separação entre este e o sorgo. O componente temporal da sua análise foi possível já que as LTRs de um mesmo elemento são idênticas no momento da inserção e, quanto mais divergentes elas são, mais tempo decorrido desde a inserção do elemento. Outro estudo, realizado por Kalendar e colaboradores (2000), apresenta um exemplo da reestruturação do genoma mediada por TEs numa população natural de cevada (*Hordeum spontaneum)*. Neste caso, a reestruturação toma a forma de uma variação no tamanho do genoma intra-específico devido a amplificação do retrotransposon com LTRs BARE-1. O número de cópias do elemento BARE-1 varia entre populações geograficamente próximas sujeitas a diferentes níveis de estresse hídrico. A correlação entre o aumento do número de cópias do retrotransposon BARE-1, o tamanho do genoma e as condições climáticas, sugere a ativação do elemento como resposta adaptativa ao ambiente. Os resultados de ambos os estudos mostraram pela primeira vez *in vivo* ou em uma população natural que os TEs podiam reestruturar um genoma.

Outro exemplo do impacto dos TEs na evolução dos genomas são os elementos de transposição em miniatura (MITEs) em duas subespécies de arroz: *japonica* e *indica*. Os MITEs são uma classe especial de elementos da Classe II não-autônomos, encontrados em um número muito elevado de cópias nos genomas, preferencialmente localizados em regiões ricas em genes. Um elemento MITE denominado mPing é ativo no genoma das subespécies de arroz. A diferença no número estimado de cópias do elemento em *japonica* e *indica* é de 70 versus 11 cópias, respectivamente. Esta diferença no número sugere uma amplificação recente desta família de MITEs, talvez durante a domesticação e possivelmente associada a diferentes condições ambientais em que cada subespécie de arroz se submeteu (Jiang *et al*. 2003, Wessler 2006).

Apesar da constatação da grande quantidade de TEs nos genomas e os ciclos de expansão observados, pouca atividade foi experimentalmente demonstrada. Alguns retrotransposons com LTRs podem ser ativados transcricionalmente sob condições de estresse biótico e abiótico. O primeiro retrotransposon ativo encontrado em plantas, Tnt1, foi isolado de fumo através da seleção de inserções no gene da nitrato redutase (NR), préviamente clonado. A família Tnt1 é o objeto de estudo deste trabalho.

A transcrição de retrotransposons com LTRs não está relacionada necessariamente a inserções novas no genoma (transposição). O ciclo de vida de um retrotranspon com LTRs inclui quatro etapas: transcrição, tradução, transcrição reversa e integração do elemento. A regulação pode acontecer em qualquer uma das etapas e como resultado limitar a taxa de transposição. A atividade dos TEs é altamente regulada no nível transcricional de maneira a estabelecer um equilíbrio que garanta a presença do elemento sem inviabilizar o genoma hospedeiro (Kumar and Bennetzen, 1999).

A variação do conteúdo de TEs entre espécies pode ter influenciado tanto a qualidade quanto a quantidade da variação genômica, e portanto impactado a evolução dos genomas. Por exemplo, quase 50 % do genoma humano é derivado de elementos L1 (LINEs) e elementos *Alu* (SINEs). Os elementos *Alu* são os elementos mais prolíficos nos genomas dos primatas, acumulando aproximadamente 1,2 milhões de cópias durante 65 milhões de anos de evolução. Os elementos *Alu* são uma fonte de instabilidade genômica (Sen *et al*. 2006). Devido a sua abundância e semelhança de seqüências, eles podem servir de sítios de recombinação resultando em deleções, duplicações e translocações (Callinan *et al.*. 2005). Os elementos *Alu* podem interromper a edição (*splicing*) normal de um gene integrando-se dentro de introns; mudar o padrão de expressão gênica, inserindo-se em regiões promotoras ou silenciando um gene através da inserção dentro do mesmo

(Deininger and Batzer, 1999). Alguns elementos *Alu* se encontram em introns de genes gerando sítios para que ocorra um *splicing* alternativo. Especificamente uns 5% dos introns de *splicing* alternativo são derivados das seqüências de *Alu,* e até tornaram-se exons; um processo que foi chamado de exonização de *Alu* (Lev-Maor *et al*. 2003).

Existem estudos que demonstram a coevolução dos elementos com os genomas hospedeiros. Casacuberta E. (2005) demonstrou que os elementos *HeT-A* e *TART* de *Drosophila melanogaster* se integram exclusivamente nas extremidades dos cromossomos, e a transposição deles é essencial para manter a integridade dos telômeros e em conseqüência a sobrevivência da espécie. A domesticação de um TE é outro exemplo de coevolução. SETMAR, um gene específico dos primatas, originou-se há 40-58 milhões de anos atrás pela fusão do domínio *histone methyltransferase* SET, com uma transposase similar ao elemento *Mariner* (Cordaux *et al*. 2006).

Os exemplos anteriores mostram as variadas formas como os TEs influenciaram a evolução dos organismos, indicando que atuaram ativamente no desenvolvimento e complexidade dos genomas.

I.5 – O retrotransposon Tnt1

Tnt1 foi o primeiro elemento de transposição caracterizado na família *Solanaceae*, sendo um dos mais bem caracterizados retrotransposons em plantas. O retroelemento Tnt1 foi isolado do genoma de fumo (*Nicotiana tabacum*) após sua inserção no gene estrutural *nitrato redutase* (NR). Para selecionar um evento tão raro Grandbastien *et al*. (1989), utilizaram protoplastos a partir dos quais podem ser regeneradas plantas. Normalmente a transcrição de Tnt1 é baixa mas fortuitamente quando a planta é exposta ao extrato de

fungo (necessário para degradar a parede celular e gerar protoplastos) é ativada. Também foi demonstrado que o retrotransposon Tnt1 pode ser ativado por vários estresses bióticos e abióticos, incluindo estresse por injúria, agentes oxidantes, infecção de patógenos e elicitores microbianos (Pouteau *et al.* 1991).

Tnt1 é um membro típico da família *Ty1/copia* de *Saccharomyces cereviseae* e *D melanogaster*, mostrando grande similaridade tanto na organização quanto em extensiva similaridade em nível de aminoácidos com esse grupo. É encontrado em elevado número de cópias no genoma de fumo (>100), e possui 5,3 kpb, sendo flanqueado por pequenas duplicações de 5 pb em seu sítio de inserção. Contém duas LTRs terminadas pelas seqüências palindrômicas 5'- TG...CA -3', comuns à grande maioria dos elementos descritos, assim como um único quadro aberto de leitura (ORF) de 1.328 aminoácidos homólogos a dois domínios codificantes de retrovírus (gag-pol) (Grandbastien *et al.* 1989). Tnt1 também contém regiões não codificantes, adjacente às LTRs: o PBS (sítio de ligação para iniciador) na extremidade 5'; e o PPT (trato de polipurinas) na extremidade 3'. Tanto o PBS quanto o PPT correspondem a sítios importantes no processo de transposição dos retroelementos com LTRs.

A expressão do elemento Tnt1 é rigidamente regulada e não é detectada na maioria dos tecidos de fumo, excetuando as raízes (Pouteau *et al*. 1991). Porém sua expressão é altamente induzida em preparos frescos de protoplastos, derivados de tecidos de mesófilo foliar de fumo; e também pode ser induzida por fatores de origem microbianos e estresses abióticos relacionados com injúria celular. Análises de expressão transiente em protoplastos de fumo mostraram que na região U3 da LTR estão à maioria dos componentes em *cis* que atuam na ativação transcricional do elemento.

O retrotransposon Retrolyc1-1, isolado de *L. peruvianum*, possui uma organização similar aos retrotransposons do grupo *Ty1/copia* e contém regiões de homologia com o elemento Tnt1 sendo considerado um membro da família Tnt1. Em *L. esculentum*, foi detectada a presença de aproximadamente 20 cópias do elemento, sendo que em *L. peruvianum* o sinal de hibridação apresentou-se bem mais forte, indicando um maior número de cópias. Esse resultado sugere que a domesticação poderia ter levado a uma perda de número de cópias homólogas a Tnt1/Retrolyc1 em *L. esculentum*. Verificou-se também que os domínios internos do elemento Tnt1 (gag, prot, endo e RT) são homólogos ao elemento Retrolyc1 de *L. peruvianum*, com exceção da região U3 presente nas LTRs do elemento, região associada com a ativação transcricional do elemento.

O retrotransposon Tnt1 é um de poucos retrotransposons de plantas com atividade transposicional demonstrada (Grandbastien *et al.* 1989). As seqüências similares a Tnt1 foram detectadas em diversas espécies em *Solanaceae*. Pelo menos três grupos principais podem ser diferenciados baseando-se na divergência da região regulatória U3, originando os grupos Tnt1A, B e C (Vernhettes *et al.* 1998; Grandbastien *et al.* 2005) (figura 5). Os elementos Retrolyc1 detectados no tomate compartilham extensas similaridades de nucleotídeo com o elemento Tnt1, exceto na região U3 (Costa *et al*. 1999; Araújo *et al.* 2001) (figura 5). A subfamília Retrolyc1, por sua vez compreende dois grupos, Retrolyc1A e B, distinguindo-se também por sua região regulatória U3 (Araújo *et al.* 2001) (figura 5).

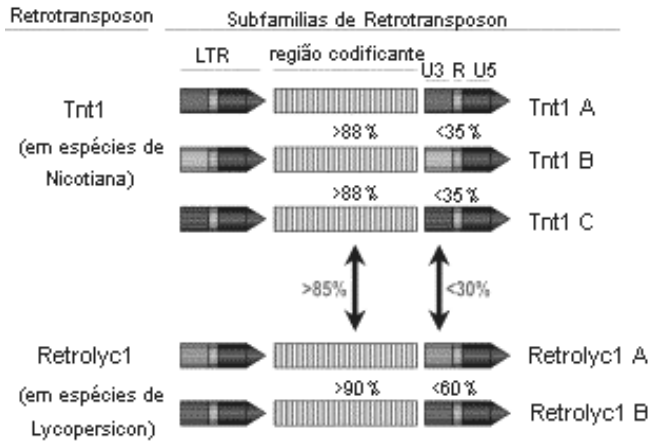

Figura 5. Comparação das regiões de Tnt1 e Retrolyc1.
Porcentagens de identidade entre as regiões U3 e regiões codificantes (baseadas no domínio da RT) entre os elementos das diferentes subfamílias Tnt1 e Retrolyc1. Adaptado de Casacuberta e Santiago 2003

I.6 - O gênero *Solanum*

A família *Solanaceae* compreende aproximadamente 90 gêneros distribuídos no mundo todo, sendo especialmente abundante nas Américas, incluindo entre 1400-1700 espécies (Mc Lean and Ivimey-Cook, 1956). Muitas espécies têm sido exploradas pelo homem desde a Antiguidade, muitas das quais foram domesticadas. Dentre as espécies com maior importância econômica, encontram-se a batata, o tomate e o fumo, que pertencem a gêneros diferentes: *Solanum*, *Lycopersicon* e *Nicotiana*, respectivamente. Com respeito ao gênero *Lycopersicon*, ainda existe controvérsia se as espécies que ele abrange pertenceriam ao gênero *Solanum,* especialmente na literatura referente ao melhoramento genético. A clasificação das espécies silvestres de tomate baseadas na morfologia as colocam dentro do um gênero deferente: *Lycopersicon* (Muller, 1940; Luckwill, 1943 e Child, 1990). Embora quando concetos biológicos das espécies são considerados *Lycopersicon* é uma subseção de *Solanum* (Rick, 1979) (para uma revisão ver Peralta e Spooner, 2000).

Dentro de *Solanum*, as batatas encontram-se no subgênero *Solanum*, seção *Petota* Dumort, subseção *Potatoe*. Todas as espécies desta seção são consideradas batatas por produzir tubérculos. Estão amplamente distribuídas nas Américas desde a região sudoeste

dos Estados Unidos (Nebraska e Colorado) até o extremo sul dos Andes. Segundo Hawkes (1990) existem 232 espécies dentro desta subseção, das quais 7 são cultivadas e 225 silvestres. As espécies cultivadas surgiram depois da domesticação, há 7000 anos atrás quando os índios começaram a coletar, armazenar e plantar os tubérculos (baseado nas análises de espécies da batata coletadas numa caverna no Perú).

Através de cruzamentos e hibridizações interespecíficas naturais entre as quatro espécies silvestres *S. acaule* (4x) *S. sparsipilum* (2x), *S. leptophyes* (2x) e *S. megistacrolubum* (2x) se originaram todas as espécies cultivadas. Elas variam em sua ploidia, resistência a fatores bióticos ou abióticos, morfologia e qualidade para consumo. Durante o cultivo da batata em tempos pré-históricos a espécie *Solanum tuberosum* (4x) se dispersou pelo mundo e transformou-se na quarta lavoura mundial, diferenciou-se em duas subespécies geograficamente muito distantes; *tuberosum* e *andígena*. A subespécie *andígena* se originou nos Andes peruanos e a subespécie *tuberosum* na ilha de Chiloé, Chile. A subespécie *tuberosum* teria originado os cultivares atualmente usados, após sua introdução na Europa (Ochoa, 1989).

As espécies silvestres possuem características de ervas daninhas adaptadas a condições ecológicas pouco favoráveis para o crescimento. Apresentam estolões, alto conteúdo de alcalóides, sementes e tubérculos adaptados a longos períodos de secas e geadas (Hawkes, 1999).

Outra subseção da seção *Petota*, *Etuberosum*, contém uma espécie, *Solanum brevidens*, que não produz tubérculos, mas que é utilizada em programas de melhoramento genético por possuir resistência a nemátóides e ser de fácil cruzamento com as batatas.

VI - CONCLUSÕES

1. O gênero *Solanum* apresenta retrotransposons da família Tnt1 constituído por três grupos de seqüênciais que se diferenciam na região U3 da LTR.
2. Analisando a região codificante da RNAseH modelos de expansão foram propostos para cada gênero a saber, *Solanum*, *Lycopersicon* e *Nicotiana*.
3. A família de retrotransposons Tnt1 amplificou-se diferencialmente entre as espécies apresentando menor número cópias em *Solanum*.
4. Uma cópia completa denominada Retrosol1 foi isolada de *Solanum oplocense* e confirma similaridade estrutural e em nível de seqüência com os membros da família Tnt1. A família de retrotransposons Tnt1 se diferencia em nível genérico através da variabilidade de sua região U3 e da região RNAseH codificante.

VII – RESUMO

Tnt1 foi o primeiro retrotransposon ativo identificado em fumo após sua inserção no gene estrutural da *nitrato redutase*. Dez anos mais tarde, um novo membro foi caracterizado em *Lycopersicon*: Retrolyc1. Neste estudo analisamos seqüências similares a Tnt1 em 20 espécies silvestres de *Solanum* e cinco cultivares de *Solanum tuberosum*. As seqüências completas foram amplificadas a partir de DNA genômico utilizando uma estratégia baseada em PCR. Os fragmentos purificados foram clonados e seqüenciados, e a análise filogenética revelou três grupos que diferem na sua região U3 [Genbank: EF620567-EF620763]. Estes fragmentos clonados foram chamados de Retrosol.

Utilizando uma abordagem de "network" com um total de 453 seqüências não redundantes e isoladas de espécies de: *Solanum* (197), *Nicotiana* (140) e (116) *Lycopersicon*, foi demonstrado que a família Tnt1 pode ser tratada como uma população para tentar resolver ramificações filogenéticas multiforcadas. A rede resultante da RNAseH revelou que as seqüências agrupam de acordo com o gênero de *Solanaceae*, sustentando uma forte associação com o genoma hospedeiro, e que a variação das seqüências na região U3 associada, caracteriza o padrão evolutivo modular dentro da família Tnt1. Dentro de cada gênero, e independentemente das espécies, quase 20% das seqüências Tnt1 analisadas são idênticas, o que indica que foram provenientes de uma cópia ativa. As relações de "network" permitiram a identificação da "seqüência mestre ou fundadora" e forneceu provas de que dentro de cada gênero essas seqüências mestre são associadas com regiões U3 distintas.

O número de cópias de seqüências similares a Tnt1 foi determinado utilizando PCR quantitativo em tempo-real. Comparando-se o número de cópias, foi revelado que em fumo as cópias de Tnt1 são abundantes e representam cerca de 2% do genoma, enquanto em tomate as cópias de Retrolyc1 são menos abundantes. As espécies de *Solanum* mostraram um número baixo de cópias de Retrosol e uma relação LTR: domínio interno de aproximadamente 2:1, sugerindo que a maioria das cópias são completas.

Utilizando uma estratégia baseada em PCR, uma seqüência completa de Retrosol de quase 5 kb foi clonada. Esta cópia possui todas as características típicas de um retrotransposon tipo *Ty1-copia* e apresenta baixa identidade de nucleotídeos na região U3 quando comparada, aos outros membros da família.

Os resultados aqui apresentados corroboram a hipótese de que a família Tnt1 estava presente logo no início da evolução de *Solanaceae*. A evidência sugere também que a região da RNAseH ficou fixada ao nível de gênero no hospedeiro, e que dentro de cada gênero a propagação foi assegurado pela diversificação da região U3.

VIII – *Abstract*

Tnt1 was the first active plant retrotransposon identified in tobacco after nitrate reductase gene disruption. Ten years later a new member was characterized in *Lycopersicon:* Retrolyc1. In this study, we performed an analysis of Tnt1-like sequences of 20 wild species of *Solanum* and five cultivars of *Solanum tuberosum.* Sequences were amplified from total genomic DNA using a PCR-based approach. Purified fragments were cloned and sequenced, and clustering analysis revealed three groups that differ in their U3 region [GenBank: EF620567-EF620763]. These cloned fragments were named Retrosol.

Using a network approach with a total of 453 non-redundant sequences isolated from *Solanum* (197), *Nicotiana* (140) and *Lycopersicon* (116) species, it was demonstrated that the Tnt1 family can be treated as a population to resolve previous phylogenetic multifurcations. The resulting RNAseH network revealed that sequences group according to the *Solanaceae* genus, supporting a strong association with the host genome, whereas tracing the U3 region sequence association characterizes the modular evolutionary pattern within the Tnt1 family. Within each genus, and irrespective of species, nearly 20% of Tnt1 sequences analyzed are identical, indicative of being part of an active copy. The network approach enabled the identification of putative "master" sequences and provided evidence that within a genus these master sequences are associated with distinct U3 regions.

Copy number of Tnt1-like sequences was determined using quantitative real-time PCR. Comparing copy number among the species revealed that Tnt1 elements are abundant in tobacco, representing around 2% of the genome while Retrolyc1 elements in tomato are less abundant. *Solanum* species show low copy number of Retrosol elements and an LTR:internal domain ratio of about 2:1, indicating that most copies are complete.

A complete sequence of Retrosol, which is nearly 5 kb long was cloned using a PCR approach. It has all the features typical of a *Ty1-copia* like retrotransposon and show low nucleotide identity in the U3 region when compared with the other members of the family.

The results presented here support the hypothesis that the Tnt1 family was present early in the evolution of *Solanaceae*. The evidence also suggests that the RNAseH region of Tnt1 became fixed at the host genus level whereas, within each genus, propagation was ensured by the diversification of the U3 region.

IX - REFERÊNCIAS BIBLIOGRÁFICAS

AMMIRAJU, J. S. S.; ZUCCOLO, A.; YU, Y.; SONG, X.; PIEGU, B.; CHEVALIER, F.; WALING, J. G.; MA, J.; TALAG, J.; BRAR, D. S.; SAN MIGUEL, P. J.; JIANG, N.; JACKSON, S. A.; PANAUD, O.; WING, R. A. **Evolutionary dynamics of an ancient retrotransposon family provides insigths into evolution of the genome size in the genus *Oryza***. The Plant Journal. doi: 10.1111/j.1365-313x.2007.03242.x. 2007.

ARAUJO, P. G.; CASACUBERTA, J. M.; COSTA, A. P. P.; HASHIMOTO, R. Y.; GRANDBASTIEN, M. A.; VAN SLUYS, M. A. **Retrolyc1 subfamilies defined by different U3 LTR regulatory regions in the *Lycopersicon* genus.** Mol Genet Genomics**.** 266:35-41. 2001.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W.; LIPMAN, D. J. **Basic local alignment search tool.** J Mol Biol. 215:403-410. 1990.

BANDELT, H. J. **Median-Joining networks for inferring intraespecific phylogenies**. Mol Biol Evol. 16:37-48. 1999.

BABUSHOK, D. V.; OSTERTAG, E.M.; KAZAZIAN, H. H. **Current topics in genome evolution: Molecular mechanism of new gene formation.** Cell.MolLife Sci.64:542-554.2007.

BENNETZEN, J. L.; KELLOG, E. A. **Do plants have a one way ticket to genomic obesity?** The Plant Cell. 9:1509-1514. 1997.

BENNETZEN, J. L.; MA, J.; DEVOS, K. M. **Mechanisms of Recent Genome Size Variation in Flowering Plants.** Annals of Botany. 95:127-132. 2005.

BOEKE, J. D.; CORCES, V. G. **Transcription and reverse transcription of retrotransposon**. Annu Rev Microbiol. 43: 403-34. 1989.

BOWEN, N. J.; JORDAN, I. K. **Transposable elements and the evolution of eukaryotic complexity**. Curr. Issues Mol. Biol. 4: 65–76. 2002.

CALLINAN, P. A.;WANG, J.; HERKE, S. W.; GARBER, R. K.; LIANG, P., BATZER, M. A. ***Alu* Retrotransposition-mediated Deletion.** J. Mol. Biol. 348:791-800. 2005.

CASACUBERTA, E.; PARDUE, M. L. **HeT-A and TART, two Drosophila retrotransposon with a bona fide role in chromosome structure for more than 60 millions years**. Cytogenet. Genome Res. 110 (1-4): 152-9. 2005.

CASACUBERTA, J. M.; VERNHETTES, S.; GRANDBASTIEN, M. A. **Sequence variability within the tobacco retrotranspsoson Tnt1 population.** EMBO J. 14:2670-2678. 1995.

CASACUBERTA, J. M.; SANTIAGO, N. **Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes**. Gene. 311: 1-11 2003.

CASACUBERTA, J. M.; GRANDBASTIEN, M. A. **Characterisation of LTR sequences involved in the protoplast specific expression of the tobacco Tnt1 retrotransposon**. Nucleic Acids Res. 11: 21(9):2087-93. 1993.

CHILD, A. **A synopsis of *Solanum* subgenus *Potatoe* (G. Don) D'Arcy [*Tuberarium* (Dun.) Bitter** (s.l.)] *Feddes Repert*. 101:209-235. 1990.

CORDAUX, R.; UDIT, S.; BATZER, M. A.; FESCHOTTE, C. **Birth of a chimeric primate gene by capture of the transposase gene from a mobile element**. Proc Natl Acad Sci U S A. 23: 103(21):8101-6. 2006.

CORDAUX, R.; HEDGES, D. J.; BATZER, M. A. **Retrotransposition of Alu elements: how many sources?** Trends in Genetics**.** 20(10):464-467. 2004.

COSTA, A. P. P.; SCORTECCI, K.; HASHIMOTO, R. Y.; ARAUJO, P. G.; GRANDBASTIEN, M. A.; VAN SLUYS, M. A. **Retrolyc-1, a member of the Tnt1 retrotransposon super-family in the Lycopersicon peruvianum genome.** Genetica. 197:65-72. 1999.

D'ARCY, W. **The Solanaceae since 1976, with a Review of its Biogreography**. Solanaceae III:Taxonomy, Chemestry, Evolution. Edited by: Hawkes, Lester, Nee, Estrada. Royal Botanic Gardens Kew and Linnean Society of London; 1991.

DEININGER, P. L.; BATZER, M. A. ***Alu* repeats and human disease**. Mol. Genet. Metab. 67:183-193. 1999.

DEININGER, P. L.; MORAN, J. V.; BATZER, M. A.; KAZAZIAN, H. H. **Mobile elements and mammalian genome evolution.** Current Opinion in Genetics & Development. 13(6):651-658.2003.

DEVOS, K. M.; BROWN, J. K. M.; BENNETZEN, J. L. **Genome Size Reduction through Illegitimate Recombination Counteracts Genome Expansion in *Arabidopsis***. Genome Research. 12:1075-1079. 2002.

DOOLITTLE, W. F.; SAPIENZA, C. Selfish **genes, the phenotype paradigm and genome evolution.** Nature 284: 601–3. 1980.

DOOLITTLE, R. F.; FENG, D. F.; JOHNSON, M. S.; MCCLURE, M. A. **Origins and evolutionary relationships of retroviruses. Q. Rev. Biol. 64**(1): 1-30. 1989.

EWING, B.; HILLIER, L.; WEDL, M.; GREEN, P: **Basecalling of automated sequencer traces using phred. I. Accuracy assessment**. Genome Research. 8:175-185. 1998.

FLAVEL, A. J.; SMITH, D. B.; KUMAR, A**: Extreme heterogeneity of Ty1-copia group retrotransposons in plants.** Mol Gen Genet. 231:233-242. 1992.

FEDOROFF, N. **Transposons and genome evolution in plants**.Proc Natl Acad Sci U S A. 20;97(13):7002-7. Review. 2000.

FESCHOTTE, C.; JIANG, N.; WESSLER, R. **Plant transposable elements: where genetics meets genomics**. Nature Reviews. 3:329-341. 2002.

FULTON, T. M.; CHUNWONGSE, J.; TANKSLEY, S. D. **Microprep protocol for extraction of DNA from tomato and other herbaceous plants.** Plant Molecular Biology Reporter. 13:207-209. 1995.

GORBUNA, V.; LEVY, A. A. **How plants make ende meet: DNA double-strand break repair**. Trends in Plant Science. 4:263-269. 1997.

GORDON, D. P.; ABAJIAN, C.; GREEN, P**. Consed: a graphical tool for sequence finishing**. Genome Research. 8:195-202. 1998.

GORDON, D.; DESMARAIS, C.; GREEN, P**. Automated finishing with Autofinish.** Genome Res . 11:614-625. 2001.

GRANDBASTIEN, M. A.; SPIELMAN, A.; CABOCHE, M. **Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics.** Nature. 26;337(6205)**:**376-80. 1989.

GRANDBASTIEN, M. A.; AUDEON, C.; BONNIVARD, E.; CASACUBERTA, J. M.; CHALHOUB, B.; COSTA, A. P. P.; LE, Q. H.; MELAYAH, D.; PETIT, M.; PONCET, C.; TAM, S. M.; VAN SLUYS, M. A.; MHIRI, C. **Stress activation and**

genomic impact of Tnt1 retrotransposons in Solanaceae. Cytogenet Genome Res. 110(1–4):229-41. 2005.

HALL, T. A. **BioEdit: a user-friendly biological sequence alignment editor and analysis program for Window 95/98/NT.** Nucl. Acids. Symp. Ser. 41:95-98. 1999.

HANAHAN, D. **Studies on transformation of E. coli with plasmids.** J. Mol. Biol. 156:557-580. 1983.

HANSON, C. N.; HESLOP-HARRISON, J. S. **Sequences and Phylogenies of Plant Pararetroviruses, Virus and Transposable Elements.** Advanced in Botanical Research. 41:165-193. 2004.

HAVECKER, E. R.; GAO, X.; VOYTAS, D. F. **The diversity of LTR retrotransposons**. Genome Biol. 5(6):225. Epub 2004 May 18. 2004.

HAWKES, J. G. **The potato: evolution, biodiversity and genetic resources**. Belhaven Press, London. 1990.

HAWKES, J. D. **The economic importance of the family *Solanaceae***. In Solanaceae IV. Edited by: Nee M, Symon DE, Lester RN, Jessop JP. Royal Botanic Gardens, Kew; 1999:1-8.

JIANG, N.; BAO, Z.; ZHANG, X.; MCCOUCH, S. R.; EDDY, S. R.; WESSLER, S. R. **An active DNA transposon in rice.** Nature 421: 163–7. 2003.

JIANG, N.; WESSLER, S. R. **Insertion preference of maize and rice miniature inverted repeat transposable elements as revealed by the analysis of nested elements**. Plant Cell. 13: 2553-2564. 2001.

JORDAN, M. A.; SAEDLER, H.; STARLINGER, P. **Stron-polar mutations in the trnasferase gene of the galactose operon E. Coli.** Mol.Gen.Genet.100:296-306. 1967.

JUKES, T. H.; CANTOR, C. R. Evolution of protein molecules. In Munro, H.N. (ed), **Mammalian Protein Metabolism.** Academic Press, New York, N.Y. pp. 21-132. 1969.

KALENDAR, R.; TANSKANEN, J.; IMMONEN, S.; NEVO, E.; SCHULMAN, A. H. **Genome evolution of wild barley (Hordeum spontaneum) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence**. Proc Natl Acad Sci U S A . 97:6603-6607. 2000.

KALMYKOVA, A. I.; GVOZDEV, V. A. **Selective expansion of the newly evolved genomic variants of retrotransposon 1731 in the Drosophila genomes**. Mol Biol Evo. 21(12):2281-9. 2004.

KASHKUSH, K.; FELDMAN, M.; LEVY, A. A. **Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat**. Nat. Genet. 33(1): 102-6. 2003.

KAZAZIAN, H. H. JR. **Mobile elements: drivers of genome evolution**. Science. 12;303(5664):1626-32. 2004.

KIDWELL, M. G.; LISCH, D. R. **Transposable elements as sources of variation in animals and plants**. Proc Natl Acad Sci U S A. 22;94(15):7704-11. 1997.

KIDWELL, M. G.; LISCH, D. R. **Transposable elements and host genome evolution**. Trends Ecol. Evol. 15:95-99. 2000.

KIMURA, M. **A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences.** J. Mol. Evol. 16: 111-120. 1980.

KUBISTA, M.; ANDRADE, J. M.; BENGTSSON, M.; FOROOTAN, A.; JONÁK, J.; LIND, K.; SINDELKA, R.; SJÖBACK, R.; SJÖGREEN, B.; STRÖMBOM, L.; STAHLBERG, A.; ZORIC, N. **The real-time polymerase chain reaction.** Mol Aspects Med. Apr-Jun;27(2-3):95-125. 2006.

KUMAR, A; BENNETZEN, J. L: **Plant Retrotransposons.** Annu Rev Genet. 33**:**479-532. 1999.

KUMAR, S.; TAMURA, K.; NEI, M. **MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment.** Briefings in Bioinformatics. 5:150-163. 2004.

LANDER, E. S, *et al*. **Initial sequencing and analysis of the human genome**. Nature. 409(6822):860-921. 2001.

LEETON, P. R. J.; SMYTH, D. R. **An abundant LINE-like element. amplified in the genome of Lilium speciosum.** Mol. Gen. Genet. 237(1-2):97-104. 1993.

LERAT, E.; BRUNET, F.; BAZIN, C.; CAPY, P**. Is the evolution of transposable elements modular?** Genetica. 107: 15-25. 1999.

LEV-MAOR, G.; SOREK, R.; SHOMRON, N.; AST, G. **The birth of an alternatively spliced exon: 3′ splice-site selection in *Alu* exons.** Science 300: 1288–91. 2003.

LIVAK, K. J.; SCHMITTGEN, T. D. **Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$Method**. Methods. 25:402-408. 2001.

LUCKWILL, L. C. **The genus *Lycopersicon*: an historical, biological, and taxonomical survey of the wild and cultivated tomatoes**. Aberdeen Univ. Stud. 120:1-44. 1943.

MA, J.; DEVOS, K. M.; BENNETZEN, J. L. **Analyses of LTR-Retrotransposon Structures Reveal Recent and Rapid Genomic DNA Loss in Rice.** Genome Research. 14:860-869. 2004.

MALIK, H. S.; EICKBUSH, T. H. **Phylogenetic Analysis of Ribonuclease H Domains Suggest a Late, Chimeric Origin of LTR Retrotransposable Elements and Retroviruses**. Genome Research**.** 11:1187-1197. 2001.

MC LEAN, R. C.; IVIMEY-COOK, W. R. **Textbook of Theorical Botany**. Vol II Longman. 1956.

MORGANTE, M. **Plant genome organisation and diversity: the year of the junk!** Current Opinion in Biotechnology. 17 (2):168-173. 2006.

MULLER, C. H. A revision of genus *Lycopersicon*. U.S.D.A. **Misc.Publ**. 382:1-28 + 10pl. 1940.

OCHOA, C. M**. Las papas de Sudamérica, Perú.** Centro Internacional da la Papa (CIP). 1989.

OGIWARA, I.; MIYA, M.; OHSHIMA, K.; OKADA, N. **Retropositional parasitism of SINEs on LINEs: identification of SINEs and LINEs in elasmobranchs.** Mol Biol Evol. 16(9):1238-50. 1999.

ORGEL, L. E.; CRICK, F. H. C. **Selfish DNA: the ultimate parasite**. Nature 284: 604–7. 1980.

PERLMAN, P.S.; BOEKE, J. D. **Ring Around the Retroelement**. Science. 303(5655): 182-184. 2004.

POUTEAU, S.; HUTTNER, E.; GRANDBASTIEN, M. A.; CABOCHE, M. **Specific expression of the tobacco Tnt1 retrotransposon in protoplasts.** EMBO J. 10(7):1911-8. 1991.

RIRIE, K. M.; RASMUSSEN, R. P.; WITTWER, C. T. **Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction**. Analytical Biochemistry. 245:154-160. 1997.

RICE, P.; LONGDEN, I.; BLEASBY, A. **EMBOSS: The European Molecular Biology Open Software Suite.** Trends in Genetics. 16(6):276-277. 2000.

RICK, C. M. **Hybridization between *Lycopersicon esculentum* and *Solanum pennellii*. Phylogenetic and cytogenetic significance**. Proc Natl Acad Sci U S A. 46:78-82. 1960.

ROZAS, J.; ROZAS, R. **DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis**. Bioinformatics. 15:174-175. 1999.

SAN MIGUEL, P.; TIKHONOV, A.; JIN, Y. K. **Nested retrotransposons in the intergenic regions of the maize genome.** Science. 274: 765-768. 1996.

SAN MIGUEL, P.; BRANDON, S.; GAUT; TIKHONOV, A.; NAKAJIMA, Y.; BENNETZEN, J. L. **The paleontology of intergene retrotransposons of maize**. Nature Genetics .20, 43 – 45. 1998.

SAN MIGUEL, P.; BENNETZEN, J. L. **Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons**. Annals of Botany *82* (Supplement A)**:**37-44. 1998.

SEN, S. K.; HAN, K.; WANG, J.; LEE, J.; WANG, H.; CALLINAN, P. A.; DYER, M.; CORDAUX, R.; LIANG, P.; BATZER, M. A: **Human genomic deletions mediated by recombination between Alu elements**. Am J Hum Genet. 79(1):41-53. 2006.

SOLEIMANI, V. D.; BAUM, B. R.; JOHNSON, D. A. **Quantification of the retrotransposon BARE-1 reveals the dynamic nature of the barley genome.** Genome. 49:389-396. 2006.

SPOONER, D. M.; PERALTA, I. E. **Classification of wild tomatoes: a review**. Kurtziana. Tomo 28 (1): 45-54. 2000.

SUONIEMI, A.; TANKSKANEN, K.; SCHULMAN, A. H. **Gypsy-like retrotransposons are widespread in the plant kingdom.** Plant J. 13(5): 699-705. 1998.

SWOFFORD, D. L. **PAUP: Phylogenetic Analysis Using Parsimony, Version 3.0 s** (Computer program and manual distributed by the Center for Biodiversity, Illinois Natural History Survey, Champaign, IL 61820) 1992.

THOMPSON, J. D.; HIGGINS, D. G.; GIBSON, T. J. **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice.** Nucleic Acids Research, 22:4673-4680. 1994.

VERNHETTES, S.; GRANDBASTIEN, M. A.; CASACUBERTA, J. M. **The Evolutionary Analysis of the Tnt1 Retrotransposon in Nicotiana Species Reveals the High Variability of its Regulatory Sequences.** Mol Biol Evol. 15(7):827-836. 1998.

VICIENT, C. M.; SUONIEMI, A.; ANAMTHAWAT-JÓNSSON, K.; TANSKANEN, J.; BEHARAV, A.; NEVO, E.; SCHULMAN, A. H.. **Retrotransposon BARE-1 and its role in genome expansion in the genus Hordeum.** Plant Cell.11:1769-1784. 1999.

VICIENT, C. M.; KALENDAR, R.; SCHULMAN, A. H: **Variability, recombination, and mosaic evolution of the barley BARE-1 retrotransposo**n. J Mol Evol. 61(3):275-91. 2005.

VOYTAS, D. F.; CUMMINGS, M. P.; KONICZNY, A.; AUSUBEL, F. M.; RODERMEL, S. R. **Copia-like retrotransposons are ubiquitous among plants.** Proc Natl Acad Sci U S A. 1;89(15):7124-8. 1992.

WESSLER, S. R. **Eukariotic Transposable Elements: Teching Old Genomes New Tricks**. Chapter 8. The Implicit Genome (Caporale L, ed.). Oxford University Press. USA. 2006.

WHITE, S. E.; HABERA, L. F.; WESSLER, S. R**. Retrotransposons in the flanking regions of normal plant genes: A role for copia-like elements in the evolution of gene structure and expression.** Proc Natl Acad Sci U S A . 91: 11792-11796. 1994.

WIKSTRÖM, N.; SAVOLAINEN, V.; CHASE, M. W. **Evolution of the angiosperm calibrating he family tree.** Proc. R. Soc. Lond. 268:2211-2220. 2001.

WRIGHT, S.; FINNEGAN, D. **Genome evolution: Sex and the transposable element**. Current Biology**.** 11:R296-R299. 2001.

XIONG, Y.; EICKBUSH, T. H. **Origin and evolution of retroelements based upon their reverse transcriptase sequences.** EMBO J. 9(10):3353-62. 1990.

ZACHAR, Z.; DAVIDSON, D.; GARZA, D.; BINGHAM, P. M. **A detailed developmental and structural study of the transcriptional effects of insertion of the Copia transposon into the white locus of Drosophila melanogaster.** Genetics. 111(3):495-515. 1985.

http://www.tigr.org/index.shtml.	J.Craig Venter Institute
http://www.sgn.cornell.edu/index.pl.	Sol genomics Network.
http://www.ncbi.nlm.nih.gov/BLAST/	NCBI BLAST
http://www.ebi.ac.uk/clustalw/	CLUSTAW
http://www.ub.es/dnasp/	Dnasp
http://www.fluxus-technology.com/index.htm	NETWORK
http://www.megasoftware.net/	MEGA3