# UNIVERSIDADE TÉCNICA DE LISBOA

## INSTITUTO SUPERIOR DE AGRONOMIA

## Comparative analysis of cytogenomic traits in Fagaceae species: an evolutionary approach

**"Tese apresentada neste instituto para obtenção do grau de Doutor"**

ANA TERESA MOUSINHO RESINA RIBEIRO

**ORIENTADOR: MARIA LEONOR MOTA MORAIS CECÍLIO**

**JÚRI:**

Presidente: Reitor da Universidade Técnica de Lisboa

Vogais: Doutora Kesara Anamthawat-Jónsson, professor of Plant Genetics, Faculty of Life and Environmental Sciences School of Engineering and Sciences University of Iceland;

Doutora Maria Wanda Sarujine Viegas, professora catedrática do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

Doutora Maria Helena Reis de Noronha Ribeiro de Almeida, professora associada do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

Doutor Rogério Paulo de Andrade Tenreiro, professor auxiliar da Faculdade de Ciências da Universidade de Lisboa;

Doutora Maria Leonor Mota Morais Cecílio, professora auxiliar do Instituto Superior de Agronomia da Universidade Técnica de Lisboa.

**DOUTORAMENTO EM BIOLOGIA**

LISBOA

2008

# *Dedicatória*

*Em memória do meu Pai*

*Francisco Rafael Resina Ribeiro*

O principal responsável pelo desenvolvimento do meu gosto pela Microscopia e Biologia Celular.

# *Agradecimentos (Acknowledgments)*

Muitas foram as pessoas que directa ou indirectamente estiveram envolvidas comigo neste trabalho e contribuíram para a conclusão desta tese.

Agradeço profundamente:

**Prof. Rui Malhó** por me ter indicado e apresentado o grupo de investigação da Prof. Wanda Viegas, no qual me senti, desde o primeiro momento, muito bem recebida.

**Prof. Leonor Morais Cecílio**, que me aceitou primeiro como bolseira de investigação e teve a coragem de me aceitar de novo como aluna de Doutoramento. Agradeço a sua orientação sempre presente, a sua disponibilidade para resolver assuntos práticos de falta de material (casa "Sport Billy") e especialmente a sua grande qualidade humana que foi crucial nos momentos difíceis.

**Prof. Wanda Viegas** por se ter disponibilizado, apesar do seu volumoso trabalho, a ler a tese, dando o seu valioso contributo científico, através das pertinentes críticas e sugestões.

À **"ciganada da ciência"**-colegas de grupo de investigação, entre estas pessoas incluo, algumas que, embora não presentes no período do desenvolvimento da tese, foram muito importantes na minha formação, ou pessoas que proporcionaram um ambiente alegre e descontraído onde uma pessoa tímida frequentemente se ria com eles, mas nunca deles: **Bruno** e **Lia**, colegas de início; Mestre (no sentido literal) **Magui** por me ensinar de uma maneira tão natural (maneira que eu sempre tentei seguir quando tive de explicar qualquer coisa a alguém, mas que sempre achei que fiquei aquém); inspiradora e acelerada **Manela**; **Guida** pela paciência com que ensina, **Pedro Costa-Nunes** pela boa vontade; **Xana** pelos ensinamentos científicos e histórias hilariantes; **Nuno Neves** pelo ser desconcertante que é; **Ana Caperta** pelas suas vertentes científicas e exotéricas; **Olga** pelos ensinamentos e instigação ao trabalho; **Felipe** por acreditar; **Augusta Barão**, o pilar e está tudo dito; **Márcia**, a demonstração de coragem; **Ana Luísa**, uma vencedora que consegue sempre ultrapassar os obstáculos; **Miguel Bento**, um colega brincalhão; **Filipa**; **Sofia Pereira**, **Rocheta**, **Margarida Padawon**, **Edna**, **Ana Sofia Pires**, **Nuno Pires**, **João** e **Larguinho** (os génios-gambás da "idade do gelo") e **Ritinha**.

<u>De uma maneira directa no Doutoramento tenho que agradecer:</u>

- Os ensinamentos e ajuda prestados pelos colegas **Sofia Pereira** na técnica de AFLP e programa de PCR com primers de rDNA; **Miguel Bento**, **Nuno Pires**, **Margarida Carvalho** e **Margarida Rocheta** nas clonagens. Disponibilidade na partilha do Citómetro de fluxo pela **Prof. Conceição Santos** e ajuda prestada pelo **Dr. João Loureiro**.
- O material biológico cedido pelo **Dr. Hachemi Merouani**, **CENASEF**, **Viveiros Stº Isidro** e **Engª. Isabel Silvestre**, **Prof. Carlos Abreu** e **Dr. Atsuchi Sakai**. Algumas pessoas, com as quais não tive qualquer contacto directo, mas que se dispuseram, sem qualquer tipo de restrições, a ajudarem desde logo.
- O espaço de estufa cedido pelos **viveiros do Departamento de Engenharia Florestal** que permitiu abrigar as plantas durante os períodos de Verão e a boa vontade em ajudar da **Engª. Carla Faria** que permitiu a obtenção de raízes de Faia num curto espaço de tempo.

Agradeço especialmente o apoio da minha **mãe**, **Marta Ribeiro**, a sua paciência e compreensão durante este período, pois sem a sua ajuda não teria sido possível. Obrigada **Drª. Manuela Correia**.

Agradeço a Bolsa de Doutoramento concedida pela Fundação para a Ciência e Tecnologia (FCT) com a ref. SFRH/BD/13319/2003.

# *Resumo*

A família Fagaceae possui 9 géneros distribuídos pelo Hemisfério Norte incluindo os três mais importantes géneros florestais europeus: *Quercus*, *Castanea* e *Fagus*. Neste trabalho procedeu-se à caracterização a nível citogenético, epigenético e molecular de algumas espécies com diferentes origens geográficas de modo a contribuir para a compreensão dos processos evolutivos nesta família. Os estudos cariotaxonómicos, basearam-se na análise da morfologia cromossómica, índice de assimetria cariotípica, e mapeamento dos *loci* de rDNA em várias espécies europeias e asiáticas. Esta análise revelou marcantes diferenças entre géneros, possibilitando a proposta de potenciais mecanismos evolutivos associados a rearranjos cromossómicos. Paralelamente, a análise do domínio funcional NOR, revelou a presença de um pseudosatélite apenas nos géneros *Castanea* e *Quercus*. A caracterização da topologia nuclear em células somáticas, baseada na distribuição de domínios heterocromáticos, padrões epigenéticos e na organização de sequências repetitivas específicas, revelou uma significativa diversidade entre as espécies estudadas. Enquanto que, novos padrões epigenéticos foram detectados em núcleos de pólen maduro de *Quercus suber*.

Os estudos realizados permitem concluir que o género *Fagus* se situa numa posição basal, o género *Castanea* possui alguma divergência entre espécies, e estudos em espécies asiáticas confirmam a homogeneidade de cariótipos no subgénero *Quercus*, indicando uma estabilização evolutiva durante a evolução das Fagaceas.



**Palavras-chave:** *Fagus*; *Castanea*; *Quercus*; evolução; citogenómica; epigenética

# Abstract

Fagaceae is a family with 9 genera distributed over the Northern Hemisphere including the three most important European forest genera: *Quercus*, *Castanea* and *Fagus*. Using cytogenetic, epigenetic and molecular approaches several species of those genera belonging to different geographic origins were characterized in order to understand the evolutionary processes in this family. Karyotaxonomic studies were based on the analysis of chromosome morphology, karyotype asymmetry index, and rDNA *loci* mapping in several European and Asiatic species. This analysis revealed several differences between genera, giving rise to the proposal of potential evolutionary mechanisms associated with chromosomal rearrangements. Moreover, the study of NOR functional domain revealed the presence of a pseudosatellite in only *Castanea* and *Quercus* genera. Nuclear topology characterization of somatic cells, based on heterochromatic domains distribution, epigenetic patterns and on the specific repetitive sequences organization revealed a significant diversity between studied species. And new epigenetic patterns were detected on mature pollen nuclei of *Quercus suber*.

The studies made can lead to the conclusion that *Fagus* genus is in a basal position, *Castanea* genus has some divergence between species and the study of Asiatic species confirmed the homogenization of subgenus *Quercus* karyotypes, pointing to its evolutionary stabilization in the evolution of Fagaceae.



**Key-words:** *Fagus*; *Castanea*; *Quercus*; evolution; cytogenomic; epigenetic

# *Resumo da Tese*

## *Introdução*

A família das Fagaceas é composta essencialmente por espécies arbóreas e alguns arbustos sendo actualmente reconhecidos nove géneros. *Quercus*, *Castanea* e *Fagus* sobressaem como sendo os géneros florestais dominantes nas zonas temperadas e Mediterrânicas do Hemisfério Norte com grande impacto ecológico e económico. Do ponto de vista ecológico constituem, tal como todas as espécies florestais, a principal forma de biomassa terrestre, contribuindo para a sequestração do carbono e reciclagem de nutrientes para além de terem um papel fundamental na preservação de habitats faunísticos e de flora que lhe estão associados. O seu contributo económico é patente na utilização da madeira para diversos fins consoante as espécies e dos frutos para consumo humano ou animal. *Quercus suber* L., espécie existente nos Montados ibéricos e que se desenvolve simpátricamente com *Quercus ilex* subsp. *rotundifolia* Lam., é explorado para a produção de cortiça o que o torna muito importante do ponto de vista económico pois permite, principalmente para Portugal, a manutenção de numerosos postos de trabalho e um elevado volume de exportações.

A caracterização taxonómica desta família tem sofrido várias alterações ao longo dos tempos quer a nível genérico, quer a nível de secções consoante os caracteres escolhidos. Foi a partir da classificação de Nixon (1989) que se estabeleceram os nove géneros reconhecidos até hoje na família Fagaceae: *Fagus* L., *Castanea* Mill., *Castanopsis* Spach., *Chrysolepis* Hjelmquist, *Colombobalanus* (Lozano, Hdz-C. & Henao) Nixon & Crepet, *Formanodendron* (Camus) Nixon & Crepet, *Lithocarpus* Bl., *Quercus* L., e *Trigonobalanus* Forman (Nixon, 1989), embora este último género continue a ser alvo de controvérsia (Manos *et al*., 2001, Chen *et al*., 2007). É também Nixon (1993) a estabelecer a mais recente classificação do género *Quercus*, baseando-se na análise cladística morfológica. Este autor reconhece dois subgéneros, *Quercus* e *Cyclobalanopsis*, este último exclusivo da Ásia. Dentro do subgénero *Quercus*, reconhece menos grupos infragenéricos que as restantes classificações (e.g., Camus, 1954, Schwarz, 1964, Trelease, 1924): *Lobatae* Loudon, os carvalhos vermelhos distribuídos no continente Americano; *Protobalanus* (Trelease) Schwarz, carvalhos

intermediários, presentes no Noroeste Americano e *Quercus sensu stricto,* carvalhos brancos distribuídos pela Eurásia e América ou exclusivamente Eurásia se considerarmos a subsecção *Cerris*. Dentro dos géneros de Fagaceae presentes na Europa, o subgénero *Quercus* está representado a nível mundial com cerca de 381 espécies, *Fagus* com 10 e *Castanea* com 7. *Fagus* possui o subgénero *Fagus*: distribuído pela América e Eurásia, e *Engleriana*, exclusivamente Asiática (Shen, 1992), enquanto que *Castanea* possui 3 secções: *Eucastanon*, com 5 espécies, incluindo *C. sativa* (europeia), *C. crenata* (japonesa) e *C. mollissima* (chinesa); *Balanocastanon* e *Hypocastanon*, cada uma com uma espécie (revisto em Dane *et al*., 2003; Lang *et al*., 2006).

Estudos filogenéticos baseados em análises moleculares apontam para uma divergência do género *Fagus* que cedo terá ocorrido na evolução das Fagaceas e que os géneros *Quercus* e *Castanea*, apesar de apresentarem algumas características morfológicas e estratégias de polinização diferentes (Manos *et al.,* 2001) se encontram bastante próximos (Manos *et al.,* 2001; Liu *et al*., 2004). Do mesmo modo foi estabelecido que *Fagus* e *Castanea* têm uma origem Asiática (Paffetti *et al*., 2007; Lang *et al*., 2007).

Os estudos evolutivos podem também recorrer à análise citogenómica que ao comparar cariótipos de diferentes espécies poderá detectar possíveis rearranjos cromosómicos associados a processos evolutivos. Também a comparação da topologia de diferentes genomas, e a correlação da estrutura, actividade e padrões epigenéticos de domínios funcionais fornecem importantes dados para a análise comparativa de diferentes espécies e géneros.

Os genomas são entidades com grande variabilidade e muito dinâmicas nomeadamente no que diz respeito aos vários processos biológicos como a replicação e a transcrição envolvendo um vasto conjunto de interacções moleculares entre DNA-proteínas e proteínas-proteínas.

Apesar das espécies de *Fagus, Quercus* e *Castanea* possuírem o mesmo número de cromossomas (2n=24), uma das espécies de *Trigonobalanus sensu lato* apresenta 2n=14 e outra espécie poliploide 2n=6x=42 (Hou, 1971 revisto em Chen *et al*., 2007). A existência de triplóides intra-específicos espontâneos foi detectada no género *Castanea* e *Quercus* (Jaynes, 1962; Burda e Shchepotiev, 1973; Butorina, 1993; Dzialuk *et al*., 2007) sendo também conhecida nestes dois géneros uma enorme facilidade de hibridação inter-específica (Jaynes, 1962; Rushton, 1993). Apesar da grande taxa de

hibridação inter-específica de onde podem resultar híbridos estéreis, férteis ou parcialmente férteis (Rushton, 1993) e apesar de níveis de introgressão consideráveis (Petit *et al*., 2003), as espécies de carvalhos simpátricos conseguem-se manter como entidades genéticas distintas, pondo em causa o conceito de espécie baseado no total isolamento genético da espécie (Whittemore e Schall, 1991, Muir *et al*., 2000; Curtu *et al*., 2007).

Os genomas eucariotas podem variar muito de tamanho o que está essencialmente associado a diferentes quantidades de sequências repetitivas. Nestas incluem-se os elementos transponíveis e em especial os retrotransposões LTR que foram invadindo os genomas ao longo de vários períodos aparecendo como os maiores constituintes dos genomas das plantas (Zhang e Wessler, 2004). Para além destas sequências repetitivas que se encontram geralmente dispersas em genomas grandes e médios (em relação à quantidade de DNA que possuem), existem ainda sequências organizadas em "tandem", como as sequências satélite. As sequências repetitivas são geralmente não-codificantes, mas no entanto podem ter um papel funcional como é o caso das sequências teloméricas e centroméricas. No entanto, outras sequências repetitivas englobam sequências codificantes como é o caso dos elementos transponíveis e dos genes ribossomais. As sequências repetitivas não codificantes estão sujeitas a pressões evolutivas mais fracas que os genes e por isso evoluem mais depressa, acumulando variabilidade tanto ao nível de dimensão como na sequência de nucleótidos, gerando muitas vezes variantes que são específicas da espécie ou até mesmo de um *locus* o que dá origem ao conceito de família de sequências repetitivas de um mesmo motivo repetido (Schmidt e Heslop-Harrison, 1998). Pensa-se também que as unidades de repetição não evoluem independentemente, mas de uma maneira concertada de modo que uma determinada sequência repetitiva poderá ser mais semelhante dentro da espécie do que entre espécies (revisto em Elder e Turner, 1995).

Neste trabalho os genomas dos três géneros *Fagus*, *Quercus* e *Castanea* foram caracterizados a nível citogenético, epigenético e molecular, através da análise de espécies europeias (*Fagus sylvatica* L., *Quercus suber* L., *Quercus ilex* subsp. *rotundifolia* Lam., e *Castanea sativa* Mill.) e asiáticas (*Quercus acutissima* Carruth., *Quercus serrata* Murray, *Castanea crenata* Sieb. & Zucc. e *Castanea mollissima* Bl.).

## *Variabilidade intergenérica, inter- e intraespecífica*

Os processos de divergência e especiação são frequentemente acompanhados por rearranjos nos complementos cromossómicos, sendo por isso a análise comparativa dos cariótipos uma ferramenta importante em estudos evolutivos. A estrutura morfológica dos cromossomas, baseada na posição do centrómero é por vezes insuficiente para fazer a identificação de todos os pares de homólogos, especialmente se estes tiverem uma estrutura muito semelhante ou se forem de pequena dimensão, como é o caso dos cromossomas das Fagaceas. Os genes ribossomais (45S e 5S rDNA) constituem os marcadores cromossómicos mais usados devido à sua natureza repetitiva e grande quantidade de cópias por *locus* e à conservação das suas sequências codificantes. São também usados porque ajudam a detectar alguns rearranjos cromossómicos de difícil detecção através da morfologia dos cromossomas e também polimorfismos de *loci* entre espécies e variedades (Jiang e Gill, 2006).

Os cariótipos de *Q. suber*, *C. sativa* e *F. sylvatica* foram estudados do ponto de vista morfométrico, do mapeamento dos *loci* de rDNA, e de sequências teloméricas. O mapeamento dos *loci* de rDNA foi também efectuado em várias espécies europeias e asiáticas de *Quercus* e *Castanea*. Todas as espécies estudadas apresentavam 2n=24 cromossomas, tendo porém sido detectado um individuo de *Q. suber* triplóide com 36 cromossomas e cujo cariótipo não apresentava rearranjos cromossómicos em relação aos cariótipos diplóides. A presença de indivíduos triplóides é normalmente associada à produção de gâmetas não reduzidos (Butorina, 1993; Zhang *et al*., 2007), no entanto, através de estudos comparativos de dimensão do grão de pólen e determinação da ploidia do núcleo generativo através de FISH, não foi possível detectar gâmetas duplicados.

Entre os três géneros estudados, *C. sativa* apresentou o cariótipo mais simétrico, isto é, com predominância de cromossomas com a posição do centrómero na região mediana e *Q. suber* apresentou o cariótipo mais assimétrico com maior heterogeneidade de classes e com maior número de cromossomas com o centrómero na região subterminal. No entanto, *Q. suber* e *C. sativa* apresentam o mesmo número e posição relativa de genes ribossomais, assim como acontece com as restantes espécies europeias e asiáticas de ambos os géneros, com excepção de *C. mollissima. F. sylvatica* apresenta o dobro dos *loci* e com diferente disposição cromossómica no que respeita aos *loci* de 45S rDNA. As sequências teloméricas foram detectadas apenas nas extremidades dos

cromossomas, embora a sua intensidade seja diferente consoante as espécies, sugerindo que o maior número de cópias detectado em *C. sativa* esteja correlacionado com a maior dimensão do genoma desta espécie em relação às outras espécies. Por outro lado, a presença de cromossomas e braços cromossómicos em *F. sylvatica* e em *Q. suber* sem sinal indica que o número de cópias varia entre os braços dos cromossomas de uma mesma espécie.

As diferenças cromossómicas detectadas nas espécies dos diferentes géneros de Fagaceas estudados possibilitaram a formulação de modelos explicativos de potenciais rearranjos cromossómicos que tenham presidido à transição de um cariótipo ancestral para o cariótipo de *Quercus* assim como à transição do cariótipo de *C. crenata* para *C. mollissima*. Assim propõem-se que os rearranjos envolveram a diminuição do número de *loci* de rDNA, reposicionamento dos *loci* de 45S rDNA e amplificação igual ou diferencial de DNA ao longo do genoma de *Quercus* levando a alterações nos tamanhos dos braços dos cromossomas. No que respeita à transição do cariótipo das duas espécies asiáticas de *Castanea* poderá considerar-se que os rearranjos envolveram apenas 2 cromossomas, com a adição de um *locus* extra de 5S e o reposicionamento de um *locus* de 45S rDNA na transição para o cariótipo de *C. mollissima.*

## *Análise comparativa da região organizadora do nucléolo (NOR) como um domínio genómico funcional*

Os seres eucariotas possuem o seu DNA complexado com proteínas, dando origem a uma estrutura mais ou menos compacta designada cromatina. A compactação processa-se por várias etapas começando com a associação do DNA a um octâmero de histonas. Estas histonas consistem em quatro tipos, H2A,H2B, H3 e H4 e possuem carácter básico devido aos aminoácidos constituintes. O complexo formado pelo DNA enrolado à volta das histonas dá origem aos nucleossomas que se encontram ligados entre si por uma porção de DNA de ligação e por uma quinta histona - H1, formando uma fibra linear polinucleossómica. A partir deste estado vão surgindo novos níveis de maior compactação até à formação do cromossoma metafásico onde se observa o nível máximo de compactação. A interacção da fibra polinucleossómica com a maquinaria da replicação e da transcrição está patente em vários modelos que afirmam que fibras de

cromatina formam ansas que saem de territórios cromossómicos interfásicos, e que entram em contacto com as diferentes maquinarias (Belmont *et al*., 1999; Blanton *et al*., 2003; van Driel *et al*., 2003; Heng *et al*., 2004; Butaye *et al*., 2004; Tetko *et al*., 2006). A cromatina pode apresentar-se em duas formas diferentes, eucromatina, quando o nível de condensação é menor em interfase do que em metafase, e heterocromatina quando o nível de condensação é muito elevado ao longo de todo o ciclo celular. A heterocromatina pode ser facultativa se só se apresentar num determinado momento do desenvolvimento, ou constitutiva quando não sofre alterações com as fases do desenvolvimento. Aos domínios eucromáticos estão associados genes potencialmente activos, enquanto que aos domínios heterocromáticos estão associadas sequências de DNA em estados transcricionalmente inertes (Turner, 2001). Num genoma para além da informação genética existe a informação epigenética que se sobrepõe à primeira e que se pode manifestar diferencialmente nos diferentes domínios de cromatina. Aos diferentes domínios de cromatina estão associados diferentes propriedades bioquímicas que constituem marcas epigenéticas associadas à manutenção da actividade génica ou ao silenciamento. Estas propriedades incluem níveis distintos de metilação de DNA e modificações pós-tradução das histonas, tais como metilações e acetilações nos aminoácidos das caudas das histonas (Tariq e Paszkowski, 2004).

Num genoma nem todos os genes, apesar de funcionais, estão activos, como acontece com os genes ribossomais. Nas plantas estes genes são compostos por centenas a milhares de sequências repetidas de rDNA 45S presentes num ou mais *loci* designados NORs. As sequências encontram-se organizadas sempre no mesmo sentido e cada unidade de repetição é composta pelo cistrão que contém os genes 18S, 5.8S e 25S e os separadores internos (ITS1 e ITS2), e pelo separador intergénico (IGS) que pode ser dividido numa zona não transcrita onde se encontram várias subrepetições, cujo tamanho e número pode ser variável, dando origem aos polimorfismos muitas vezes encontrados nestes IGS; e numa parte transcrita, o separador transcrito externo (ETS) (Moss e Stefanovsky, 1995; Pikaard, 2000). Os NORs activos dão origem aos nucléolos interfásicos e às constrições secundárias nos cromossomas metafásicos que podem ser detectados pelo nitrato de prata devido à manutenção de proteínas relacionadas com a maquinaria transcricional nesses locais (Goodpasture e Bloom, 1975; Roussel *et al*., 1996; Zurita *et al*., 1998). Devido ao grande número de cistrões, os NORs activos apresentam domínios condensados e descondensados que estão relacionados com sequências de genes não transcritas e transcritas, respectivamente (revisto em Neves *et*

*al*., 2005a) e que se encontram associados a diferentes marcas epigenéticas. Em *Arabidopsis suecica* verificou-se que a metilação da histona H3 na lisina 9 (H3K9me2) se encontra co-localizada com NORs inactivos, enquanto que NORs activos acumulavam domínios ricos em H3K4me3, a marca epigenética associada à transcrição activa (Lawrence *et al*., 2004). Estes estudos de regulação de transcrição de genes ribossomais em plantas estão até certo ponto de acordo com o modelo apresentado em mamíferos (revisto em Preuss e Pikaard, 2007). Nos mamíferos o complexo remodelador nucleolar - NoRC - medeia o silenciamento dos genes ribossomais ao recrutar desacetilases de histonas. Estas enzimas retiram grupos acetil dos aminoácidos das histonas que se encontram nos promotores, tornando mais fortes as ligações entre DNA e histonas e permitindo que se estabeleçam marcas para o recrutamento de metiltransferases para a lisina na posição 9 da histona H3 que por sua vez parece ter uma grande interacção com a metilação de DNA. Finalmente o recrutamento de DNA metiltransferases permite a metilação *de novo* dos genes ribossomais, essencial para o silenciamento destes genes (Santoro e Grummt, 2005).

O estudo da estrutura do *locus* do NOR activo nos géneros *Fagus, Quercus* e *Castanea* revelou a existência de uma estrutura cromossómica pouco condensada de forma esférica e de tamanho variável na extremidade dos cromossomas nucleolares composta por sequências ribossomais 45S. Esta estrutura foi designada pseudosatélite e está presente em todas as espécies de *Quercus* e *Castanea* estudadas, excepto em *C. mollissima* onde se encontra um satélite típico. FISH com sondas de rDNA 45S mostram que se podem detectar três domínios nos NORs interfásicos activos de *Quercus* e *Castanea:* (i) um domínio proximal condensado de onde emanam (ii) um ou dois fios de cromatina intranucleolar que termina (iii) num bloco condensado à periferia do nucléolo e que corresponde ao pseudosatélite. Os pseudosatélites homólogos podem ser heteromórficos e encontrarem-se individualizados ou fundidos com a mesma orientação ou com orientações opostas.

A detecção simultânea de marcas epigenéticas como a metilação de citosinas (DNA) e H3K9me2 e FISH com sondas de rDNA 45S revelaram que apenas um *locus* de rDNA 45S está activo em *F. sylvatica*, o que está de acordo com a presença de apenas um nucléolo nas células meristemáticas desta espécie. A metilação do DNA é particularmente evidente nos pseudosatélites de *Castanea sativa* não havendo diferenças entre pseudosatélites heteromórficos presentes na mesma célula.

A caracterização molecular dos NORs dos três géneros demonstrou que nas sequências repetitivas de rDNA 45S estão presentes vários rearranjos estruturais para além das unidades canónicas de repetição ordenadas com a mesma orientação. Estes rearranjos incluem regiões de ETS truncadas por uma sequência palindrómica do gene 18S, o que leva à presença de unidades ribossomais mais pequenas. Essas unidades estão presentes nos três géneros com dimensões e homologias diferentes: são maiores em *C. sativa* e *Q. ilex* subsp. *rotundifolia* e menores em *F. sylvatica* e possuem 91% de homologia entre *C. sativa* e *Q. ilex* subsp. *rotundifolia* e apenas 74% de homologia com *F. sylvatica*. Além disso, em *C. sativa* foi também detectada a inserção de uma região pequena do gene mitocôndrial de rDNA 26S em substituição do separador intergénico IGS.

Os mecanismos propostos para o primeiro tipo de rearranjos são ciclos de quebra durante translocações assimétricas recíprocas em anafase e fusão durante a replicação. Estes ciclos podem levar a heteromorfismos evidenciados citológicamente nos pseudosatélites e formação de sequências palindrómicas através da inversão de sequências; ou recombinação somática durante o período em que os dois pseudosatélites se encontram fundidos em direcções opostas.

O mecanismo que tem sido apontado como o mais comum para a inserção de fracções de DNA de organelos no genoma nuclear em vários organismos é a reparação de quebras nas duas cadeias de DNA através da junção terminal não homóloga que depende de pouca (caso da inserção em *C. sativa*) ou nenhuma homologia entre as extremidades.

## *Topologia da cromatina nuclear e distribuição de marcas epigenéticas no núcleo interfásico*

A arquitectura do núcleo interfásico depende da natureza e distribuição de sequências de DNA que estão relacionadas com os diferentes domínios de cromatina e com a expressão génica e que são controlados e mantidos por processos epigenéticos.

A disposição e organização dos cromossomas no núcleo interfásico tem sido considerada dependente do tamanho do genoma, do estado de desenvolvimento e do tipo de diferenciação celular (Cowan *et al*., 2001). Genomas grandes como os dos

cereais possuem uma orientação Rabl (Rabl, 1805) com os cromossomas dispostos linearmente e com os telómeros e centrómeros localizados em pólos opostos. Genomas pequenos como o de *Arabidopsis thaliana* não possuem essa orientação (Dong e Jiang, 1998) embora em cada cromossoma, os centrómeros e os telómeros apresentem territórios distintos estando estes últimos dispostos à volta do nucléolo (Fransz *et al*., 2002; Fang e Spector, 2005). Parece também haver um arranjo conservado dentro do núcleo eucariota, onde porções de cromossomas ricos em genes se encontram mais ao centro e domínios cromossómicos com poucos genes se encontram mais frequentemente à periferia do núcleo (Cremer *et al*., 2001; Tanabe *et al*., 2002)

A organização dos domínios eucromáticos e heterocromáticos e as respectivas marcas epigenéticas, estão também consideradas como dependentes do tamanho do genoma (Houben *et al*., 2003). Em plantas com genomas pequenos a maioria das sequências repetitivas, tanto em série (com a excepção de genes ribossomais e sequência teloméricas) como dispersas (retrotransposões) localizam-se preferencialmente em regiões centroméricas, formando zonas condensadas visíveis ao microscópio, designadas cromocentros de onde emanam os domínios eucromáticos mais ricos em genes (Fransz *et al*., 2002). Em plantas com genomas grandes a distinção entre domínios heterocromáticos e eucromáticos pode ser menos evidente devido essencialmente à dispersão dos elementos transponíveis (Houben *et al*., 2003), embora se possam reconhecer domínios heterocromáticos em genomas grandes localizados subtelomericamente (Vershinin *et al*., 1995) ou intersticialmente (Siljak-Yakovlev *et al*., 2002).

A caracterização de padrões epigenéticos e mapeamento de sequências repetitivas específicas permitiram a análise comparativa da organização genómica nos géneros *Fagus, Quercus* e *Castanea*.

A organização do tipo cromocentro observou-se em *F. sylvatica* e em *Q. serrata* enquanto que nas restantes espécies de *Quercus* e *Castanea* se detectou um tipo de organização sem cromocentros, e portanto, semelhante ao descrito para espécies de plantas com genomas grandes. Apesar da presença notória de cromocentros em *Q. serrata,* observou-se também uma tendência para a organização em Rabl. Em *F. sylvatica* as marcas epigenéticas associadas aos domínios heterocromáticos não se limitam aos cromocentros mas estão também presentes na restante cromatina. Os diferentes tipos de organização interfásica encontrada em espécies de *Quercus* (que possuem o mesmo número de cromossomas e tamanho de genomas semelhantes) sugere

que a organização do genoma em termos de domínios eucromáticos e heterocromáticos é independente do seu tamanho, mas estará dependente da natureza e distribuição das sequências repetitivas que o compõem.

O genoma de *Q. ilex* subsp. *rotundifolia* foi estudado do ponto de vista molecular tendo sido isoladas e caracterizadas, tanto molecular como citogeneticamente, cinco sequências repetitivas. Estas cinco sequências estão representadas de um modo diferente nos três géneros estudados: estão todas ausentes do genoma de *F. sylvatica;* uma das sequências fazendo parte de uma sequência tipo retrotransposão LTR, está presente tanto em espécies europeias como asiáticas do género *Quercus* e em *C. sativa;* uma terceira sequência demonstrou ser específica do género *Quercus* e as restantes são específicas de *Q. ilex* subsp. *rotundifolia.* O padrão de distribuição destas sequências nos núcleos interfásicos revelou-se disperso excepto numa das sequências específica de *Q. ilex* subsp. *rotundifolia* que possui um padrão mais agrupado.

## *Marcas epigenéticas no grão de pólen maduro de Quercus suber L. (Fagaceae)*

A organização da cromatina e as marcas epigenéticas que lhes estão associadas são dinâmicas e dependentes da fase do ciclo celular, da fase do desenvolvimento e da diferenciação celular. A acetilação da lisina nas histonas H3 e H4 tem um carácter reversível dependente da actividade de acetiltransferases que adicionam grupos acetil a partir de acetil-coenzima A e de desacetilases, que os retiram (Kuo e Allis, 1998), estando fortemente associadas às fases do ciclo celular (Jasencakova *et al*., 2000; 2001; 2003). A metilação do DNA está dependente de três metiltransferases distintas consoante o contexto e pode ficar ausente após a replicação através da não manutenção na replicação ou ser retirada através de enzimas desmetilases (Agius *et al*., 2006; Morales-Ruiz *et al.*, 2006). Um baixo nível global de metilação de DNA em vertebrados e em plantas pode ter um efeito marcante no desenvolvimento dos organismos sugerindo um importante papel desta marca epigenética na regulação da expressão génica quer directamente ou através do seu efeito na manutenção de padrões específicos da estrutura da cromatina (revisto em Finnegan *et al.*, 1998, 2000). No entanto, existem

processos de desenvolvimento que necessitam da desmetilação do DNA como o que associa a expressão de genes activos no endosperma à sua desmetilação diferencial (Gehring *et al*., 2006). Também a metilação das histonas tem uma natureza reversível pois recentemente foram isoladas desmetilases de histonas que se demonstraram essenciais no desenvolvimento e diferenciação dos organismos (revisto em Benevolenskaya, 2007).

A caracterização dos padrões epigenéticos em células diferenciadas não somáticas foi realizada em núcleos das células generativas e vegetativas do grão de pólen maduro de *Quercus suber*. As marcas epigenéticas associadas à transcrição e ao silenciamento de genes, encontram-se uniformemente mais representadas nos núcleos das células generativas do que nos núcleos das células vegetativas. Esta dualidade de marcas epigenéticas presente no núcleo generativo está de acordo com as evidências de transcrição, nomeadamente de genes codificantes de proteínas envolvidas na divisão celular (Okada *et al*., 2007). O núcleo vegetativo possui menor quantidade de marcas epigenéticas associadas ao silenciamento (5-mC e H3K9me2) do que à transnscrição (H3K4me3, H4TA), ambas com um padrão de distribuição (ponteado) semelhante à observada em células somáticas, embora mais disperso. Este padrão mais disperso poderá estar associado ao relaxamento da cromatina decorrente de um menor teor da histona de ligação H1, no núcleo da célula vegetativa (Tanaka *et al*., 1998). O baixo nível de metilação de DNA presente nestes núcleos contrasta com o padrão observado em espécies de herbáceas (Oakeley *et al*., 1997; Janousek *et al*., 2000; Janousek *et al*., 2002) e poderá ser relacionado com a fertilização retardada presente no sobreiro pois nesta espécie decorrem cerca de três meses desde a polinização até à fecundação (Boavida *et al*., 1999). Assim, o nível de alterações epigenéticas, nomeadamente as relacionadas com o silenciamento génico, parece ser determinantes para assegurar as estratégias de fertilização adoptadas por cada espécie.

## *Conclusões*

Os estudos realizados permitiram uma compreensão mais profunda sobre a evolução dos genomas de espécies dos três géneros de maior relevância ecológica e económica na família Fagaceae e contribuíram para o conhecimento da história evolutiva desta família, possibilitando as seguintes conclusões:

-*Fagus* é dos géneros estudados o que ocupa a posição mais basal e *Quercus* e *Castanea* estão evolutivamente próximos o que foi concluído através do mapeamento físico de *loci* de genes ribossomais; da presença de uma estrutura cromossómica, aqui designada como pseudosatélite, comum a *Quercus* e *Castanea*, mas ausente em *Fagus;* de uma grande homologia entre algumas sequências repetitivas como é o caso da sequência do tipo retrotransposão LTR (Rot 10) ou de separadores de genes ribossomais 45S, mais propriamente de regiões ETS, entre *Quercus* e *Castanea*. Durante o decurso do trabalho surgiram resultados que contradizem visões estabelecidas, como a orientação ordenada apenas num sentido de unidades de repetição completas de rDNA 45S somente posta em causa por Caburet *et al*. (2005) ou ainda a relação entre a organização interfásica dos genomas e a sua quantidade de DNA e as organização das marcas epigenéticas que lhes estão associadas.

-O género *Quercus* subgénero *Quercus,* apesar de possuir bastante mais espécies que o género *Castanea*, parece ter estabilizado evolutivamente a nível de cariótipo enquanto que o género *Castanea*, com apenas 5 espécies na secção *Eucastanon*, apresenta marcadas divergências nas três espécies estudadas. A análise de espécies asiáticas de *Quercus* (*Q. acutissima* e *Q. serrata*) foi essencial para confirmar a conservação de cariótipos de *Quercus* subgénero *Quercus* por todo o Hemisfério Norte. Por outro lado, propomos que a migração da secção *Eucastanon* do Japão para China e depois para a Europa (Lang *et al*., 2007) levou a que ocorresse no castanheiro europeu *C. sativa*, um efeito epigenético essencialmente sob a forma de uma acentuada metilação de DNA no pseudosatélite do NOR major, e uma restruturação cariotípica do castanheiro chinês *C. mollissima*. Uma vez que *C.sativa* apresenta um cariótipo semelhante ao do castanheiro japonês, este rearranjo em *C. mollissima* deve ter ocorrido após a divergência entre *C. mollisima* e *C. sativa*.

Este trabalho permitiu ainda inferir a possibilidade de existência de paleopoliploidia durante a evolução da família Fagaceae através do conhecimento de espécies consideradas relíquias com outros números de cromossomas e a possibilidade de ter havido uma duplicação dos genomas ancestrais de onde pode ter resultado o dobro dos *loci* ribossomais no género basal Fagus.

## *Perspectivas futuras*

Paralelamente, considera-se que seria interessante desenvolverem-se futuramente estudos que permitissem uma melhor compreensão sobre:

- Os processos evolutivos associados à emergência dos diferentes cariótipos presentes na família Fagaceae, através do mapeamento de *loci* de rDNA em espécies com 2n=14 como *Trigonobalanus doichangensis;* com 2n=42 como o hexaploide *Trigonobalanus verticillata,* e espécies do género *Chrysolepsis*.

- A hipótese de uma paleopoliploidia ancestral na família Fagaceae, detectando simultaneamente o período em que ocorreu a duplicação através da análise de conjuntos de genes únicos de ESTs, tal como anteriormente utilizado no detecção de paleopoliploidia no género Populus por Sterck (2005) ou em outras plantas modelo por Blank and Wolfe (2004).

- As estratégias adoptadas pelas plantas que apresentam o processo de fertilização retardada, através da análise comparativa do transcriptoma das células vegetativas e generativas do grão de pólen de *Q. suber,* e estabelecer uma correlação com as marcas epigenéticas que estes núcleos possuem.

# *List of Abbreviations*

| | |
|---|---|
| AFLP | amplified fragment length polymorphism |
| Ag-staining | stainining with silver nitrate |
| AI | asymmetry index |
| α– | anti-, antibody against |
| anti-dig-FITC | anti-digoxigenin conjugated fluorochrome fluorescein isothiocyanate, emission in the green region of the light spectrum |
| APTES | 3-aminopropyltriethoxy-silane |
| BLASTn | BLAST program algorithm which search a nucleotide database using a nucleotide query |
| BLASTx | BLAST program algorithm which search protein database using a translated nucleotide query |
| bp, kb, Mb | base pair(s), kilobase, megabase |
| BSA | bovine serum albumin |
| cDNA | copy DNA |
| CI | centromeric index |
| DAPI | 4',6-diamidino-2-phenylindol |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytidine triphosphate |
| DECON | detergent |
| dGTP | deoxyguanosine triphosphate |
| DNA | deoxyribonucleic acid |
| DNase | deoxyribonuclease, enzyme degrading DNA |
| dNTP | nucleotides mixture |
| DSBs | double-stranded breaks |
| dTTP | deoxythimidine triphosphate |
| dUTP | deoxyuracyl triphosphate |
| *Eco*R I | restriction enzyme isolated from some strains of *Escherichia coli.* It is an endonuclease which creates sticky ends with 5' end overhangs |
| EDTA | ethylenediaminetetracetic acid |
| EST | expressed sequence tag |
| ETS | external transcribed spacer which makes part of the IGS of the 45S rDNA unit |
| FAA | formaldeyde acetic acid fixative |
| FISH | fluorescence *in situ* hybridization |
| *g* | gravitational acceleration (relative centrifuge force) |
| g, mg, μg, ng | gram, milligram, microgram, nanogram |
| h(rs), min, sec | hour(s), minutes, seconds |
| H3K4me3 | trimethylation of the lysine 4 of histone H3 aminoacids tail |
| H3K9me2 | dimethylation of lysine 9 of histone H3 aminoacids tail |
| H3K9Ac | acetylation of lysine 9 of histone H3 aminoacids tail |
| H4K5,K8,K12, K16Ac | acetylation of lysines in the positions 4,5,8,12 and 16 of the aminoacids tail of histone H4 |
| H4TA | the same as H4K5,K8,K12,K16Ac |
| IGS | intergenic spacer between 26S and 18S of the 45S rDNA unit |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| ITS1,2 | internal transcribed spacers 1,2 of the 45S rDNA unit |
| l, ml, μl | litres, millilitres, microlitres |
| LB | Luria Broth medium |
| LTR | long terminal repeat of a retrotransposon |
| μm, nm | micrometer, nanometer |
| M, mM, μM | molar, millimolar, micromolar |
| M, m | centromere at median point, centromere at median region |
| *matK* | plastidial maturase K gene |
| mya | million years ago |
| *Mse* I | restriction enzyme isolated from *Micrococcus species.* It is an endonuclease which creates sticky ends with 5' end overhangs |
| NHEJ | nonhomologous end-joining |
| NOR | nucleolar organizer region |

| | |
|---|---|
| NoRC | nucleolar remodelling complex |
| NUMT | nuclear mitochondrial DNA |
| PBS | phosphate buffered saline |
| PCR | polymerase chain reaction |
| rDNA | ribosomal DNA. Consists of two classes: 5S rRNA genes and non-transcribed spacer, and 45S rRNA genes and intergenic spacers. |
| RNA | ribonucleic acid |
| RNA pol I | RNA polymerase I |
| RNase A | ribonuclease, enzyme degrading RNA |
| RNase H | enzyme degrading RNA in RNA-DNA hybrids |
| rpm | revolutions per minute |
| rRNA | ribosomal RNA |
| sm | centromere at submedian region |
| spp | species |
| SC | secondary constrinction |
| SDS | sodium dodecyl sulphate |
| siRNA | small interfering RNA |
| SSC | saline sodium citrate |
| SSR | simple sequence repeat |
| st | centromere at subterminal region |
| streptavidin - Cy3 | streptavidin conjugated to fluorochrome cyanine 3, emission in the red region of the light spectrum |
| subsp | subspecies |
| Taq | thermostable DNA polymerase from the thermophilic bacterium *Thermus aquaticus* |
| TBS | Tris buffered saline |
| TE | Tris-EDTA buffer |
| Tris = Trizma Base | tris[hydroxymethyl] amino-methane |
| v/v | volume per volume |
| w/v | weight per volume |
| WPB | nuclear isolation buffer for woody plants |
| X-Gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |
| 5-mC | 5-methylcytosine; methylation of cytosines at the 5 position of the pyrimidine ring |

# Prologue

Fagaceae is a vast family of arboreal plants with around 1000 known species but only 9 genera (Manos *et al.*, 2001). The major contribute to this high number of species belongs to the *Quercus* genus. In this genus, there are 531 known species (Govaerts and Frodin, 1998), about half of the entire number species in this family. Three of the most important forest genera, at the ecological and economic level, in the North Hemisphere are *Quercus*, *Castanea* and *Fagus* where they are dominant trees in temperate forests or Mediterranean ecosystems. Forest trees are the principal form of terrestrial biomass, they provide environmental benefits such as carbon sequestration, nutrient cycling as well as forest ecosystems play a major role in the preservation of biological diversity. The large size of trees enables the development of physical and chemical environments that influences their own performance and that of interacting organisms such as mutualists (e.g., pollinators, seed dispersers, herbivore predators) and antagonists (e.g., herbivores and pathogens). Some tree species are deciduous while others are evergreens and the size of the leaves also varies, which can also contribute to the specificity of the associated fauna and flora. Besides this ecological aspect, various trees species are also of enormous economic significance as a source of high quality timber, plywood and cork, and fruit or seed production for consumption by humans and as animal feed. For example, oaks have a remarkable economic importance in wine producer's countries as barrels of oak wood are essential for wine maturation, and for wine to achieve its characteristics (Pérez-Prieto *et al*., 2002). The evergreen oak woodlands (*montado* in Portugal) cover an area of about 2–2.5 million ha in the Iberian Peninsula (Castro *et al*., 1998) being the dominant tree species the *Quercus suber* L.(cork oak) and the sympatric subspecies the *Quercus ilex* L. subsp. *rotundifolia* Lam. (holm oak). They are mainly located in the southern regions and are exploited for agroforestry, contribute to preserve some habitats of ecological importance and also stand out for their resistance to fire (reviewed in Costa and Pereira, 2007). These woodlands where cork is extracted are responsible for the maintenance of a large amount of employment, and markedly accounts for the Portuguese exports (reviewed in Pereira, 2007; Mendes, 2007). The holm oak has also an important economic role since its acorns are used for natural

feeding the black Mediterranean pigs, an important traditional product (reviewed in Carvalho, 2007). Portugal is also one of the most important producers of *Castanea sativa* Mill. (sweet chestnuts) for human consumption (Cortizo *et al*., 1996) and several cultivars are mainly cultivated in the Northeast of Portugal (Gomes-Laranjo *et al*., 2005). *C. sativa* is also explored for its wood quality and its many applications like furniture, flooring, woodcraft, among other products, and it has been an essential component for the economy and culture of rural population (reviewed in Paiva, 2007). The genus *Fagus* is also represented in Portugal by *Fagus sylvatica* L (European beech), although it is not an autochthones species, we can find it in forest parks and botany gardens with several ornamental cultivars, like for instance *Fagus sylvatica* Purpurea Group. *Fagus* is the genus that named the family and it is the most abundant broadleaved forest tree in Europe and western Asia. *F. sylvatica* is well distributed throughout all Europe, especially in North and Central Europe, where its timber is well explored, being used for parquet flooring, wood pavement and bentwood furniture, and also charcoal as its calorific power is extremely good comparing to other timbers. *F. sylvatica* fruits give excellent rations to animals due to their high content of fat matter (Chalupa, 1986).

## *Aims of this work*

Trees are not only overexploited but also understudied in many aspects although their tremendous economic and ecological value namely in Mediterranean ecosystems. Their size and life span makes them difficult subjects for experimental investigations (Linhart, 1999 in Petit and Hampe, 2006). From an evolutionary point of view the Fagaceae family have been studied first by morphological characters and then by cladistic analysis based on molecular sequences data, originating controversies and contradictions about the systematic of this family. Besides genetics also epigenetic systems are essential for genomic function controlling genome structure and integrity, and contributing to plant evolution (Grant-Downton and Dickinson, 2006). Apart from two comparative molecular cytogenetic studies in some European and American species of *Quercus* subgenus *Quercus* (Zoldoš *et al*., 1998; Zoldoš *et al*., 1999), there is lack of knowledge in other members of this important family.

The fundamental aim of this work is to deep our knowledge about the three European represented genera of the Fagaceae family. This study will include a cytogenomic approach with genetic and epigenetic characterizations. By comparing traits and studying evolutionary processes we were able to characterize species and understand their evolutionary position among the Fagaceae family. Therefore we intend to contribute to a deeper knowledge into Fagaceae species with high ecological and economic interest.

# *I. Introduction*

## I.1. Systematics of the Fagaceae family

Fagaceae is a vast family with around 1000 known species, making it one of the largest families within the order of Fagales (Manos *et al*., 2001). The number of species has been added and subtracted along the years, making its classification very controversial, especially when different characters are considered, even in the same morphological register. Throughout the times, the classification of the Fagaceae family suffered many alterations, including sometimes only the sectional name of some genus, and also different names for the same species. In this work we will adopt the most recent classifications for the Fagaceae family which are summarized in Table I.1. This family includes nine currently recognized genera, exclusively distributed in the Northern Hemisphere: *Fagus* L., *Castanea* Mill., *Castanopsis* Spach., *Chrysolepis* Hjelmquist, *Colombobalanus* (Lozano, Hdz-C. & Henao) Nixon & Crepet, *Formanodendron* (Camus) Nixon & Crepet, *Lithocarpus* Bl., *Quercus* L., and *Trigonobalanus* Forman (Nixon, 1989; Manos *et al*., 2001; Li *et al*., 2004). The Southern Hemisphere genus *Nothofagus* which was once placed amongst the Fagaceae family by sharing a number of common characteristics, such as cupule fruit structure with *Fagus* (Forman, 1964; Hutchinson, 1967), is now placed in its own family, Nothofagaceae (Nixon, 1989). However, some controversial still persists concerning three genera of the Fagaceae family. Manos *et al*. (2001) recommended the inclusion of *Colombobalanus* and *Formanodendron* in the same genus of *Trigonobalanus* on the basis of combined phylogenetic analysis. This has been already established by the morphological based classification of Forman (1964) and Lozano *et al*. (1979). Phytogeographic, cytological and morphological characters (Chen *et al.,* 2007) also reject the classification of three genera proposed by Nixon and Crepet (1989). Combined analysis of DNA sequences from six regions of three plant genomes (plastid, mitochondrial, and nuclear) (Li *et al*., 2004) yielded parsimony trees very identical to the ones from ITS and *matK* combined analyses (Manos *et al.*, 2001) (Fig. I.1),

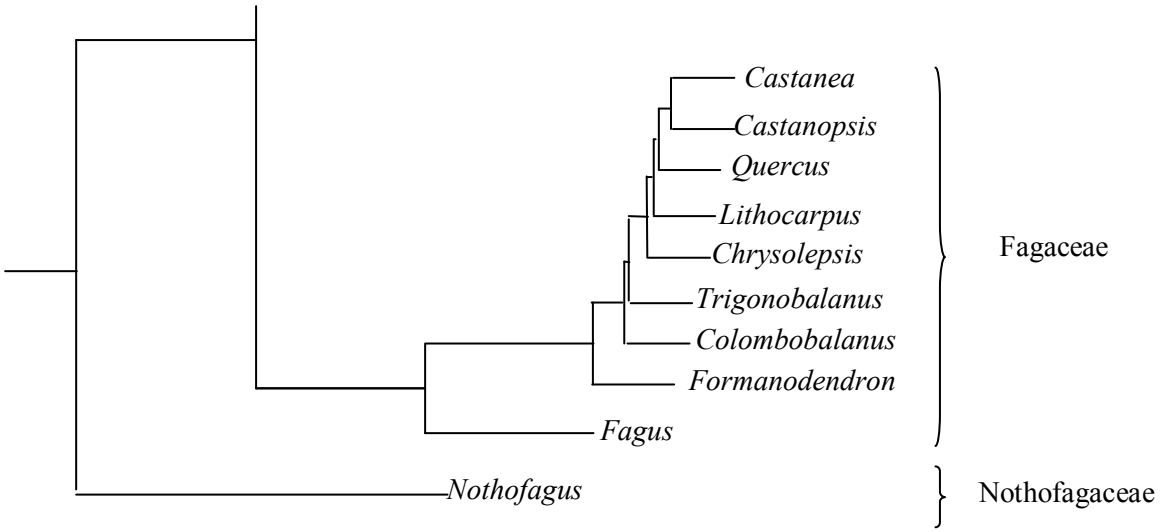



Fig. I.1 – Proposed phylogenetic relationships of Fagaceae genera adapted from Manos *et al*. (2001) and Li *et al*. (2004).

The divergence of *Fagus* DNA sequences when comparing with those of other family members suggests that, although basal within the family, it may represent a quite derived lineage. This seems somehow curious since this genus is the basis of the name of this entirely family. *Quercus* and *Castanea* were also brought closer in these two studies although sharing some common morphological features like hypogeal cotyledons and fruit wall anatomy (Soepadmo, 1968), they have different flower cupules and pollination strategies (wind *vs* insect, respectively) (reviewed in Manos *et al*., 2001).

**Table I.1**- Systematics of the Fagaceae family with some examples of species

| Genus | Subgenus | Section | Group | Nº Species | Example | Distribution |
|---|---|---|---|---|---|---|
| *Fagus* L. | *Fagus* Shen | | | **7** | *Fagus sylvatica* L | Eurasia and America |
| | *Engleriana* Shen | | | **3** | *Fagus okomotoi* Shen | Asia |
| *Formanodendron* (Camus) Nixon & Crepet | | | | **1** | *Formanodendron doichangensis* (Camus) Nixon & Crepet | Asia |
| *Colombobalanus* (Lozano, Hdz-C. & Henao) Nixon & Crepet | | | | **1** | *Colombobalanus excelsea* (Lozano, Hdz-C. & Henao) Nixon & Crepet | South America |
| *Trigonobalanus* Forman | | | | **1** | *Trigonobalanus verticillata* Forman | Asia |
| *Chrysolepsis* Hjelmquist | | | | **2** | *Chrysolepis chrysophylla* (Douglas ex Hook.) Hjelm.<br>*Chrysolepis sempervirens* (Kellog) Hjelmquist | North America |
| *Lithocarpus* Blume | *Pasania* Camus | | | **300** | *Lithocarpus hancei* (Benth.) Rehder<br>*Lithocarpus densiflorus* (Hook. & Arn.) Reh. | Asia<br>North America |
| | *Cyclobalanus* Camus | | | | *Lithocarpus hatsumii* Soepadmo | Asia |
| | *Lithocarpus* Camus | | | | *Lithocarpus ruminatus* Soepadmo | |
| | *Synaedris* Camus | | | | *Lithocarpus pulcher* (King) Markgr. | |
| | *Gymnobalanus* Camus | | | | *Lithocarpus havilandii* (Stapf) Barnett | |
| | *Nova* Manos | | | | *Lithocarpus truncatus* (King ex Hook.f.) Rehder & E. Wilson | |

| | | | | | | |
|---|---|---|---|---|---|---|
| *Quercus* L. | *Quercus* (Camus) Nixon | *Lobatae* Loudon | | **381** | *Q. rubra* L | North and South America |
| | | *Protobalanus* (Trelease) Schwarz | | | *Q. chrysolepis* Liebm | western North America |
| | | *Quercus* Nixon | *sensu stricto* | | *Q. petraea* (Matt.) Liebl.<br>*Q. robur* L.,<br>*Q. virginiana* Ten. | Eurasia and America |
| | | | *Cerris* | | *Q. suber* L<br>*Q. ilex* L.<br>*Q. acutissima* Carruth<br>*Q. serrata* Murray | Eurasia |
| | *Cyclobalanopsis* (Oersted) Schneider | | | **150** | *Quercus. sessilifolia* (Blume) Schottky | Asia |
| *Castanopsis* Spach | | | | **120** | *Castanopsis rockii* Camus | Asia |
| | | | *Fissa* Barnett | | *Castanopsis fissa* (Champ. ex Benth.) Rehder & Wilson | |
| *Castanea* Miller | | *Eucastanon* Camus | | **5** | *C. mollissima* Bl.<br>*C. seguinii* Dode<br>*C. crenata* Sieb. & Zucc.<br>*C. dentata* (Marsh.) Borkh<br>*C. sativa* Mill | Eurasia and North America |
| | | *Balanocastanon* Camus | | **1** | *C. pumila* Mill | North America |
| | | *Hypocastanon* Camus | | **1** | *C. henryi* | Asia |

### I.1.1. Fagus, Quercus and Castanea

*Fagus* L is a small genus with only 10 species distributed through Europe, Asia and North America (Shen, 1992 reviewed in Denk, 2003). On the basis of distinct morphological features, the genus *Fagus* L was divided by Shen (1992) in two subgenera: *Engleriana* with only three species exclusively from East Asia; and *Fagus* with the remaining species occurring predominantly in Eurasia and one species from North America. This subgenus includes *Fagus sylvatica* L, distributed from south-western Asia to Europe (Shen, 1992 reviewed in Denk, 2003).

Several taxonomic criteria have been produced for genus *Quercus* based on morphological characters (e.g., Camus, 1954; Nixon, 1993; Schwarz, 1964 and Trelease, 1924). These criteria also show the taxonomic controversies mainly due to the intraspecific morphological variations that may be produced by natural hybridisation (Burger, 1975) and adaptation to ecological changes in the environment (Van Valen, 1976). As a result, the classification at the subgeneric, sectional, and subsectional levels is still uncertain, and so is the taxonomic ranking of some taxa. The most current classification of *Quercus* genus (Nixon, 1993) was the first to be based on explicit morphological cladistic analysis and recognized fewer infrageneric groups than previous classifications. It broadly followed Camus infrageneric classification, but utilized Loudon's sectional names (Loudon, 1838 reviewed in Nixon, 1993), trying to make a typification of sectional names. This classification recognized two subgenera, *Quercus* and *Cyclobalanopsis* (Oersted) Schneider. The latter is restricted to the Southeast Asia and is distinguished by several flower characters, such as expanded stigmatic surfaces on the pistillate flowers and small, inconspicuous bracts which subtend single-staminate flowers, like in *Q. sessilifolia* (Blume) Schottky. Based on 17 morphological characters, as for instance, the position of abortive ovules, the persistence of leaves, and the acorn maturation timing, the subgenus *Quercus* is divided into three sections: *Lobatae* Loudon, the red oaks, distributed in the North, like *Q. rubra* L., and South America; *Protobalanus* (Trelease) Schwarz, the golden cup or intermediate oaks, distributed in the western North America, like *Q. chrysolepis* Liebm.; and *Quercus sensu stritctu* (Manos *et al*., 1999), the white oaks, widespread over North and Central America and Eurasia, e.g., *Q. alba* L, *Q. robur* L., *Q. virginiana* Ten. Two groups of white oaks are sometimes recognized as subsections, *Ilex* and *Cerris*, but

since the exact relationships of these groups are uncertain it is usually considered only the *Cerris* group (Nixon, 1993; Manos *et al.*, 1999; Manos *et al*., 2001) represented for example, by the European species *Q. suber* L. and *Q. ilex* L. and the Asian species *Q. acutissima* Carruth. and *Q. serrata* Murray (Manos *et al*., 2001; Borgardt and Nixon, 2003). Due to its enormous phenotypic variability, there has been a lot of controversy concerning the classification of *Q. ilex* L as a single species. Some authors considered *Q. ilex* as having two distinct subspecies: *Quercus ilex* subsp. *ilex* and *Quercus ilex* subsp. *rotundifolia*, the last with two possible classifications [= *Quercus ilex* L. subsp. *ballota* (Desf.) Sampaio; *Q. ilex* L- subsp. *rotundifolia* (Lam.) O. Schwarz *ex* Taborda de Morais] while others have recognized the existence of two independent species: *Quercus rotundifolia* vs *Quercus ilex*. These two independent species or subspecies differ in the shape of the leaf and taste of the fruit as well as in their geographical distribution. The first one is strictly Mediterranean, it occurs in the Iberian Peninsula and North of Africa, while the other one does not occur spontaneously in Portugal, is sub-mediterranean, present in some parts of Spain and in the European coast from France to Greece (reviewed in Capelo and Catry, 2007).

The genus *Castanea* Miller is divided into three sections mainly on the basis of reproductive structures (Jones, 1986) and comprises seven species (Johnson, 1988). Section *Eucastanon*, characterized by the presence of three nuts per cupule, contains five species widely distributed over the temperate zone of the Northern Hemisphere and where is included *C. mollissima* Bl. and *C. seguinii* Dode from China, *C. crenata* Sieb. & Zucc. (Japanese chestnut), both occurring in eastern Asia. One species is native to eastern North America, *C. dentata* (Marsh.) Borkh. (American chestnut). This species, once one of the dominant trees, has been almost wiped out by the chestnut blight disease. The remaining species is *C. sativa* Mill. (European sweetchestnut), which is native to southern Europe and western Asia. The Section *Balanocastanon*, characterized by one small nut per cupule that opens in two halves, has a single species *C. pumila* Mill. (American chinkapin) with two varieties distributed in the southeastern United States. Section *Hypocastanon*, characterized by a single nut per cupule contains also only one species, *C. henryi* (Skan) Rehder & Wilson (Chinese chinkapin) and is found in a restricted area in southeast China (reviewed in Dane *et al*., 2003; Lang *et al*., 2006).

## I.2. Genetic Variability within the Fagaceae family

To date, the genome size of species within Fagaceae's family has only been estimated for *Fagus sylvatica* (Gallois *et al*., 1999), *Castanea sativa* (Barow and Meister, 2002) and seven species of *Quercus* (Olszewska and Osiecka, 1984; Bennet and Smith, 1991; Favre and Brown, 1996; Zoldoš *et al*., 1998). Genome sizes range from 490 Mb/1C and 544 Mb/1C in *Quercus sessilis* and *Fagus sylvatica*, respectively, to 941 Mb/1C and 980 Mb/1C in *Castanea sativa* and *Quercus ilex*, respectively. The interspecific genome variation measured among seven oak species by Zoldoš *et al*. (1998) revealed only a 6% variation with a small variability found between two populations of *Q. petraea* and no variation between four populations of *Q. robur*. In addition, no clear intraspecific genome size variation was found when analyzing four varieties of *Fagus sylvatica* (Gallois *et al*., 1999).

Although the chromosome numbers are not known in all members of the Fagaceae family, all species of genera *Castanea*, *Castanopsis*, *Fagus, Lithocarpus* and *Quercus* studied show 2n=24 chromosomes, (Darlington and Wylie, 1955; Jaynes, 1962; Gallois *et al*., 1999; Zoldoš *et al*., 1999; Chokchaichamnankit *et al*., 2008), being the exceptions, *Trigonobalanus doichangensis* (Camus) Forman, also known as *Formanodendron doichangensis* according to Nixon and Crepet (1989) classification, with 2n=14 (Chen *et al*., 2007) and *Trigonobalanus verticillata* (Forman) a hexaploid with 2n=6x=42 (Hou, 1971 reviewed in Chen *et al*., 2007). Among angiosperm trees, polyploidy, the heritable condition of possessing more than two equal (autopolyploid) or different (allopolyploid) complete sets of chromosomes is common (Bennett *et al*., 1982; Clausen *et al*., 1982; Barral *et al*., 1995; Anamthawat-Jónsson and Thórsson, 2003; Barcaccia *et al*., 2003). In the Fagaceae family, few spontaneous triploids were reported in the genus *Castanea* (Jaynes, 1962) and in *Q. robur* (Burda and Shchepotiev, 1973; Butorina, 1993) and *Q. petraea* (Dzialuk *et al*., 2007). Frequency of triploids in a mixed stand of *Q petraea* and *Q robur* was around 0.48% (Dzialuk *et al*., 2007). Although triploid oaks are capable of producing acorns, they have rather low fertility, due to meiosis disturbances (Butorina, 1993). On the other hand, triploid trees have other advantages, as compared with their diploid progenitors, since triploids seem to be more tolerant to drought, and more resistant to pathogens and pests probably due to enhanced production of various secondary plant metabolites (Levin, 1983; Sherald *et*

*al*., 1994). Triploid oaks differ from other individuals of similar age because they are unusually large, having all of the characteristics of elite trees (Butorina, 1993). Spontaneous autopolyploidy is not exclusively of Fagaceae family as it has been observed in several other tree species including *Acacia dealbata* (Blakesley *et al*., 2002) or *Ulmus americana* (Sherald *et al*., 1994).

### *I.2.1. Interspecific hybridization*

Natural interspecific hybridization is known to be a common process among trees of *Castanea* and *Quercus* genera (Jaynes, 1962; Rushton, 1993). The most documented events of natural hybridization are within the genus *Quercus* L. represented by hundreds of species with extensive hybridization (Rushton, 1993). Hybridization occurs even between oaks that are morphologically and physiologically very different as a result of poor reproductive barriers. Morphological intermediacy is the major and often the only criterion used in assessing the status of putative oak hybrids. Characters are usually restricted to leaf, fruiting structures or buds and bark (reviewed in Rushton, 1993). Inference of hybridization based on morphological characters, especially in oaks which possess a wide intraspecific variability, remains however limited and can lead to wrong conclusions (Curtu *et al*., 2007). The comparative analysis of DNA sequences can also assess levels of hybridization in natural populations. Through these molecular techniques it was established that there is significant gene flow between sympatric *Quercus* species without apparent morphological intermediacy (Rushton, 1993; Bacilieri *et al*., 1996; Valbuena-Carabana *et al*., 2005). Although high level of hybridization that leads to considerable introgression, sympatric oak species are able to remain distinct, so that species concepts that rely on total genetic isolation between species to explain their distinctness clearly are not applicable to *Quercus* (Whittemore and Schall, 1991, Muir *et al*., 2000; Curtu *et al*., 2007). Astonishingly, oaks hybrids may be completely sterile, fertile, or present an intermediate behaviour (Rushton, 1993), This variability in fertility may result from compatibility of pollen-pistil interactions, from the competitive ability of pollen genotypes, or on the overlapping of geographical, phenological and ecological factors. It was shown that the cross between the two sympatric oaks *Q. ilex* subsp. *rotundifolia* and *Q. suber* was only possible in one direction and that the lack of seed set observed in the cross ♀ *Q. suber* x ♂ *Q. ilex*

subsp. *rotundifolia* is due to the inability of pollen tubes to penetrate the transmitting tissue after germination (Boavida *et al*., 2001). Partial incompatibilities among certain interspecific crosses were also seen in *Castanea* genus (Jaynes, 1961 review in Jaynes, 1962). Some hybrid combinations are promising tools to use in breeding programs directed to improve wood and fruit properties, as to enhance stress tolerance and disease resistance. *Castanea sativa* for instance produces high quality timber and fruit but is susceptible to the ink and blight diseases induced by fungi. A resistance to these diseases can be obtained by crossing the susceptible *C. sativa* with the resistant Asiatic species *(C. mollissima* or *C. crenata)* as male progenitor (Valdiviesso and da Costa, 2006). Since *C. mollissima* or *C. crenata* do not have desirable timber and growth rate, strategies involving backcrosses for a number of generations offer a way to incorporate disease resistance and timber quality in the same trees.

## I.3. Eukaryotic genome organization

### I.3.1. Repetitive sequences as major constituents of plant genomes

Eukariotyc genomes vary greatly in size, i.e., in the amount of DNA they carry. Although variation of haploid DNA contents (C-values) cannot be correlated with either the complexity of the organism or number of genes they carry, which constitutes the so called C-value paradox (Gall, 1981), it has a great effect on how genomes are organized in the interphase nucleus. In Angiosperms some variation in genome size is caused by differences in gene number, especially due to recurrent episodes of polyploid formation or segmental duplication, followed by varying degrees of gene loss (Bennetzen, 2002). Most of the variation in genome size in plants is however caused by differences in the amounts of repetitive DNA (Flavell *et al*., 1974).

Repetitive sequences are a characteristic of all eukaryotic genomes and represent between 50 and 90% or more of all DNA (Heslop-Harrison, 2000). Repetitive sequences are DNA motifs that can range in length from a single base to thousands of bases, repeated many hundreds or thousands of times. These sequences can be organized tandem in the genome, in which one copy follows the other in an array of many tens or even thousands of copies, or can be dispersed in the genome (Schmidt and

Heslop-Harrison, 1998). Tandem repeats comprise a significant portion of the repetitive DNA and include satellite families, micro- and minisatellites, which classification is based on the repeat unit length and array size. Therefore, the microsatellites have 2-5 bp repeats and array size of 10-100 bp, the minisatellites have 6-100 bp and array size of 0,5-30 kb, and satellite DNA have a variable repeat unit length (being the most common 140-180 bp and 300-360 bp) with arrays spanning up to 100 Mb (reviewed in Sharma and Raina, 2005). Satellite families are non-coding sequences but some satellite DNA corresponds to functional sequences as telomeric or centromeric repeats, essential to stabilize chromosome ends or enable chromosome segregation during mitosis and meiosis, respectively. But repetitive sequences can also code for genes like in ribosomal genes, which are also organized tandem in the genome, or in transposable elements. Individual sequences of a particular repeated motif may vary in both their copy number and exact sequence, giving rise to the concept of sequence families. This is the case, for instance, of the multigene 5S rRNA family which is represented in the *Triticeae* by two classes differing in non transcribed spacers' length and nucleotide sequences and mapped to different chromosomes (Dvorák *et al*., 1989; Reddy and Appels, 1989; Mukai *et al*., 1990).

Repetitive non-coding DNA has a rapid evolution when compared to genes because it is under weaker selective pressures (Charlesworth *et al*., 1994). They may acquire large-scale variations in their sequence and copy number over an evolutionary time-scale (Schmidt and Heslop-Harrison, 1998), which results in species-specific repeat variants and/or novel sequence families, and in the DNA content of eukaryotes spanning several orders of magnitude. Also, repeating units within different repetitive sequences evolve non-independently through concerted evolution, resulting in a sequence similarity of repeating units that is greater within than among species (reviewed in Elder and Turner, 1995).

There are two major classes of transposable elements: Class I, the retroelements (RNA transposable elements) and class II, transposons (DNA transposable elements), although both are mobile DNAs with different families, they differ in structure and in the way they make the transposition (reviewed in Bennetzen, 2000a). However, most of the transposable elements are retroelements, particularly the long terminal repeat LTR kind, that invaded genomes in different periods, and are now one of their most abundant components (Zhang and Wessler, 2004). LTR-retrotransposons comprise 60% or more of many large plant genomes like maize, wheat and barley, but less than 50% of the

small rice genome and around 10% of the smaller Arabidopsis genome (reviewed in Bennetzen, 2002). It has often been suggested that most dispersed repeats are likely to be, or have evolved from, transposable elements (Flavell 1986a; Smyth, 1991), in particular from LTR retrotransposons such as *Ty1-copia*-like or *Ty3-gypsy*-like elements. However, there are also dispersed sequences which are not related to retrotransposon-like sequences such as *Hch 1* of wild barley (Hueros *et al*., 1993) or *pBO3* of rice (Kiefer-Meyer *et al*., 1996). Many transposons sequences are also very difficult to identify because usually they are only identified through the presence of a transposase-like protein and the majority of sequences are non-autonomous, deletion derivatives, without transposases (Wicker *et al*., 2003).

### *I.3.2. Nuclear chromatin organization*

In order to compact the large amount of DNA, eukaryotes complex DNA with proteins, forming the chromatin, which is packaged in a separate compartment of the cell called the nucleus. The first order of DNA packing is the 11 nm linear nucleosome fibre: the DNA wind around the surface of an octamer of core histones (two of each histones H2A, H2B, H3 and H4) and each nucleosome is connected with its neighbours by a short segment of more accessible linker DNA. This polynucleosome string is further folded into a 30 nm chromatin fiber with the participation of the linker histone H1 associated with the linker DNA, giving rise to a zigzag two-start helix arrangement model (a double helical structure) and former known as the solenoid model structure (arranged in one helix) (review in Tremethick, 2007). Less is known about the way in which these fibres are further packed within the nucleus to form the highest-order structures leading to large chromatin domains and chromosomes territories, but several models have been proposed for nuclear structures above the 30 nm chromatin fibers and differential compaction in different chromosomes regions involving cis-acting sequences like boundary elements that interact with promoters, *Locus* Control Regions (LCRs) and Scaffold/Matrix Attachment Regions (S/MARs) (Belmont *et al*., 1999). DNA sequences known as boundary or insulator are found in organisms ranging from yeast to humans and prevent interactions between enhancers and promoters when located between them, setting up independent domains of gene regulation. Evidence for a role of insulators in the establishment of chromatin domains has come in part from the

analysis of the *gypsy* insulator of *Drosophila* (review in Labrador and Corces, 2002). Recently, it was suggested that the insulator-bounded domains form loops in the nucleus (Blanton *et al*., 2003). LCRs, have been found to regulate gene activity at the level of the cluster. (review in van Driel *et al*., 2003). The most cited model proposes that the chromatin fibres form loops anchored by S/MARs which bind to the proteinacious nuclear matrix (Avramova *et al*., 1998; Michalowski *et al*., 1999; Hancock, 2000; Tikhonov *et al*., 2000; Nickerson, 2001; Jackson, 2003). S/MARs are AT-rich DNA sequences which vary in length from a few hundred to several thousand bp, and are found in all higher eukaryotes (Bode *et al*., 1996). It has been suggested that the proposed loop organization of the chromatin triggered by the S/MARs corresponds to functional domains involved in gene regulation, based on the observation that plant transgene expression is highly stabilized by flanking inserted genes with S/MARs (Butaye *et al*., 2004). Furthermore, the recent results on a genome scale are supportive of an important role of S/MARs in spatio-temporal transcriptional regulation in Arabidopsis. They indicate that the presence of S/MARs within introns is the dominating mechanism for S/MAR mediated tissue-, organ-, and development-dependent transcriptional regulation in plants (Tetko *et al*., 2006).

FISH analysis and recent studies in which specific *loci* are fluorescent labeled in living cells show that the three-dimensional structure of a chromosome inside the nucleus is dynamic and is affected by its transcriptional status. Moreover, chromatin loops occurs in every genomes although their disposition may vary a lot whether they are large or small genomes (review in van Driel and Franz, 2004).

Within the interphase nucleus the disposition and organization of chromosomes is dependent on genome size, developmental stage and cell differentiation (Cowan *et al*., 2001). Large genomes like *Triticum aestivum* with 16979 Mb/1C (Bennett and Smith, 1976) have an interphase Rabl orientation (Rabl, 1805) with the chromosomes runnig string-like through the nucleus with telomeres and centromeres at opposite poles (Abranches *et al*., 1998). Small genomes like *Arabidopsis thaliana* 157Mb/C (Bennett *et al*., 2003) do not have a Rabl orientation, nonetheless telomeres and centromeres have defined territories and individual chromosomes occupy discrete domains within the nucleus (Fransz *et al*., 2002; Fang and Spector, 2005). However, there seems to be no specific order in the arrangement of chromosomes inside the eukaryotic nucleus, except that gene-dense chromosomes more often reside in the centre, whereas gene-

poor chromosomes reside more frequently in the periphery (Cremer *et al*., 2001). This arrangement is evolutionary conserved (Tanabe *et al*., 2002).

## I.4. Epigenetics marks and chromatin dynamics

### I.4.1. Euchromatin and Heterochromatin

Two types of chromatin have been described to make part of most eukaryotic genomes: euchromatin and heterochromatin. Cytologycally, euchromatin is defined as the genome fraction which is less condensed in metaphase and decondensed in interphase, presumably allowing for transcription. While heterochromatin is a term initially proposed by Heitz (1928) to describe the fraction that is cytologically more compact than euchromatin, that maintains high levels of condensation and remains deeply staining (heteropycnotyc) throughout the cell cycle. Heterochromatin can be constitutive, when the high degree of condensation is permanent or facultative when the compaction is developmentally regulated as a result of cellular differentiation, and can be of any size, from a gene promotor to a whole genome (Craig, 2005), being the best-studied example, the inactive X chromosome in female mammals (Turner, 2001).

Other features refine the characterization of these fractions at the molecular level. Euchromatin is associated with genome regions that replicate throughout S phase, with normal frequency of meiotic recombination, mainly composed of unique DNA sequences and with high density of genes. While heterochromatin is associated with genome regions that replicate late in the S phase, with low frequency of meiotic recombination, mainly composed of repetitive DNA sequences and with low gene density, thus being a domain largely transcriptionally inert. The structure of both types of chromatin is also different: heterocromatin is less accessible to nucleases due to the regularly spaced nucleosomes, while euchromatin has irregular nucleosomes distribution (reviewed in Richards and Elgin, 2002). However, there are some exceptions: the characteristic late replication of heterochromatic regions first described by Lima-de-Faria and Jaworska (1986) and well established for a wide range of organisms (reviewed in Gilbert, 2002), does not apply to the heterochromatic centromeres and silent mating type cassettes of *Schizosaccharomyces pombe* (Kim *et*

*al*., 2003). In fact, it was suggested that heterochromatin is a term that describes many different types of relatively condensed chromatin, perhaps with different features and roles. One example is the α- and β-heterochromatin of *Drosophila*, the first type is rich in satellite DNA and apparently lacks genes and the second type is less condensed, deficient in tandem arrays and contains transposable elements interspersed with actively transcribed genes (reviewed in Bennetzen, 2000b).

Apart from the level of condensation and the genomic regions associated, euchromatin and heterochromatin have distinct biochemical properties that are thought to provide epigenetic marks for the formation of higher order structure and for maintaining gene activity and/or gene silencing. These properties include distinct levels of DNA methylation, and diverse covalent post-translational histone modifications associated with non-histone proteins (Tariq and Paszkowski, 2004). All these histone modifications as well as DNA methylation provide extra information present in the eukaryotic nucleus besides the one contained in the DNA molecule. This information can be seen as epi-information since is on the top of genetic information and can even control it being therefore entitled its transmission as epigenetics. Chromatin structure is mostly important to several biological processes, including transcription. The organization and composition of chromatin can control the access of regulatory transcription factors and RNA polymerase to DNA. The 30 nm chromatin fibre needs to unfold to the 11 nm polynucleossome string to generate a template for transcription, which requires specific post-translational histone modifications (Zhang and Reinberg, 2001).

### *I.4.2. Histone posttranslational modifications*

Histones are basic proteins due to their high content (10-20%) of basic aminoacids (lysine and arginine) and therefore they establish strong although reversible bounds to the DNA molecule which has acidic residues due to its phosphoric acid molecules. Consequently, at physiological pH they have positive and negative net charges, respectively

Canonical histones (H2A, H2B, H3 and H4) can have variants, non-allelic isoforms that show sequence variations, like for instance, CenH3 or H3.3, with distinct biological functions, which can be incorporated throughout the cell cycle, unlike the

canonical histones that are incorporated during DNA replication (Henikoff and Ahmad, 2005).

Both variants and canonical core histones can assume an extreme dynamical role in chromatin states and biological functions. The first ones by the deposition at particular sites in the genome that can be permanent like CenH3, which is exclusively localized at the centromere, and forms part of the centromere complex (Talbert *et al*., 2000) or can be transitory like H3.3 or tissue-specific (Fernandez-Capetillo *et al*., 2003); and canonical histones can be subjected to a myriad of posttranslational modifications, due to multiple modifications of histone N-terminal tails that can comprise acetylation, phosphorylation, methylation, ubiquitylation, glycosylation, ADP-ribosylation, carbonylation, and sumoylation in different aminoacids (Fuchs *et al*., 2006).

Amongst these modifications, the acetylation and methylation of selected lysine residues of histones H3 and H4 seem to be particularly involved in the regulation of transcription (Richards and Elgin, 2002; Horn and Peterson, 2006). In this way, it has been proposed an "histone code" model in which the sequence of individual or combined modifications states of lysine on histone tails, including acetylation and methylation provide signals for recruitment of appropriate chromatin remodelling factors, which in turn alter chromatin states and affect transcriptional regulation (Jenuwein and Allis, 2001; Richards and Elgin, 2002). Also the histones variants like H2A.Z and H3.3, and linkers histones can be considered accessory factors in transcriptional regulation (Craig, 2005). Besides being a mark of active chromatin, H3.3 is often enriched in histone modifications associated with transcriptional activity (Ahmad and Henikoff, 2002; Loyola *et al*., 2006; Sun *et al*., 2007).

#### *I.4.2.1. Histone acetylation*

It is known for a long time that the addition of acetyl groups, using acetyl-coenzyme A as a cofactor, into aminoacid tails of histones is reversible and is catalized by histone acetyltransferases (HATs) and histone deacetylases (HDs, HDAs, or HDACs), respectively (Kuo and Allis, 1998). In plants, the N-terminal of lysines in positions 5, 8, 12, 16, 20 of H4 and 9, 14, 18, 23 of H3 are acetylatable by HATs and deacetylatable by HDACs (reviewed in Fuchs *et al*., 2006).

At the gene level, acetylation of histones has been correlated with transcriptional activity (Tian *et al*., 2005). The genes residing in euchromatin fraction are usually

associated with hyperacetylation of several lysine residues in histones H3 and H4 while the repressed heterochromatic state is associated with hypoacetylation in histone H4 (Berger, 2007). In transcriptional active chromatin, the instability of nucleosomes and access to DNA can be achieved by the weakening of interactions between histones and the DNA by acetylation of lysine residues which neutralizes the positive charges of the histones tails and thereby decreases their affinity for negatively charged DNA. This loosens of the nucleosome structure and facilitates the access and binding of transcription factors to genes transcribed by RNA polymerase II or III (Lee *et al*., 1993; Vettese-Dadey *et al*., 1996). However at the chromosomal or subnuclear level the modulation of acetylation of histone H4 seem to be dependent of the cell cycle and correlated with differential timing replication of euchromatin and heterchromatin (Jasencakova *et al*., 2000; 2001; 2003). At higher levels of organization, histone acetylation of H4, in particular, has been related in euchromatic and heterochromatic domains with replication and/or post-replication-associated processes such as repair and the maintenance of DNA methylation, and chromatin remodelling (Fuchs *et al*., 2006).

#### *I.4.2.2. Histone methylation*

Histones can be methylated on arginine (R) or lysine (K) aminoacids present in their N-terminal tails. Protein arginine methyltransferases (PRMTs) catalyze the transfer of methyl groups from S-adenosyl-L-methionine (SAM) to the nitrogens of arginine residues, and can be either mono- or dimethylated (Gary and Clarke, 1998). The histone lysine methyltransferases (HKMTs) do not appear to have a consensus binding domain for the cofactor SAM, however they are characterized by the presence of one invariant protein motif: the SET domain that is present in different groups of proteins involved in gene silencing or activation, like *Drosophila* PEV-modifier SU(VAR)3-9, *Polycomb*-group protein Enhancer of zeste [E(Z)] and the *trithorax*-group protein *Trithorax* (TRX). Thus, the methylation of histone lysines is made by several specific HKMTs, which vary not only with the lysine residue involved but also with the eukaryotic organism (reviewed in Jenuwein, 2001 and in Sims *et al*., 2003).

For a long time, histone methylation was seen as an irreversible condition, However, two classes of lysine demethylases (HKDMs) have recently been identified, one of which is the jumonji class, which removes methyl groups from H3K4me2 and me3, but also from H3K9me2 and me3 (reviewed in Berger, 2007).

In contrast to acetylation, methylation of lysines does not change the electric charge of the amino group. Histone lysine methylation marks are associated with either heterochromatin or euchromatin depending on the position of the methyl group but also on the level of methylation of the lysine residue (mono-, di- or tri). The different level of methylation of lysine residues can have different distributions among and between groups of eukaryotes. In plants, all methylation states of lysine 4 of histone H3 (H3K4) are restricted to euchromatin. This is also true for lysine 36 of histone H3 in Arabidopsis (H3K36). By contrast, only mono or dimethylation of lysine 9 of histone H3 (H3K9me1,2) and monomethylation of lysines 27 and 20 of histones H3 and H4 (H3K27me1 and H4K20me1) are conserved heterochromatin marks, (reviewed in Fuchs *et al*., 2006).

### *I.4.3. DNA methylation*

Another epigenetic mark implicated in transcription regulation is the methylation of cytosines on the carbon 5 of the pyrimidine ring giving rise to the modified base 5methyl-cytosine (5mC). Methyl groups are added using SAM as the cofactor, to cytosines after DNA replication by *de novo* and maintenance methyltranserase enzymes. During replication methylation on the new DNA strand is absent since DNA polymerase uses cytosine, therefore the methylation mark needs to be re-established on the newly synthesized strand using the pre-existing methyl groups on the complementary strand as a guide. In this way, the methylation state can be perpetuated indefinitely by maintenance methyltransferases that use hemi-methylated DNA as a substrate. In both plants and animals, cytosine is primarily methylated in the GpC dinucleotide context. However, in plants, cytosine is also methylated in CpNpG, and less abundantly in asymmetrical sequences context as CpHpH (where H is any nucleotide but G) (Gruenbaum *et al*., 1981; Pradhan and Adams, 1995).

Methyl groups are added to cytosine by three classes of methyltransferases: MET1 methyltransferase maintains CpG methylation by methylating newly synthesized DNA, and is required to faithfully maintain methylation patterns at repetitive and single copy sequences during both sporogenesis and gametogenesis (Kankel *et al*., 2003; Saze *et al*., 2003); CpNpG methylation is maintained by chromomethylase 3 (CMT3) that is unique to plants and, to a lesser extent by the domain rearranged methyltransferase

family (DRMs) (Bartee *et al*., 2001; Cao and Jacobsen, 2002). DRMs are responsible for establishing *de novo* methylation in all sequence contexts and appear to be directed by small RNA molecules (Chan *et al*., 2004). DNA methylation can also be a reversible process independent of DNA replication, since DNA demethylases, such as Demeter (DME) and Repressor Of Silencing 1 (ROS 1) (Agius *et al*., 2006; Morales-Ruiz *et al*., 2006) have been found.

In eukaryotes, this DNA base modification can directly interfere with transcription by preventing the binding of transcription factors that require contact with cytosine or blocking transcriptional machinery attached to promoter regions of genes by altering DNA structure (Bird and Wolffe, 1999), and indirectly by influencing nucleosome conformation and stability by its connection to several histone modifications (Richards and Elgin, 2002; Grewal and Elgin, 2002) resulting in a cascade of events involving ATP-dependent nucleosome remodellers, complexes which use the energy of ATP hydrolysis to slide nucleosomes and alter the spacing of nucleosome arrays (Becker and Horz, 2002), thus modulating the conversion between eu- and heterochromatin configurations according to the established epigenetic code (Grewal and Elgin, 2002).

In somatic cells, genes residing in the euchromatin fraction are usually associated with low levels of DNA methylation and the repression state of chromatin is associated with hypermethylation of cytosines (Berger, 2007). DNA methylation is also thought to provide defense against genome invasion of transposable elements by immobilizing these elements (Martienssen and Colot, 2001). CMT3 is the methylase involved in thid process since it is preferentially targeted to transposons as shown by genome-wide methylated DNA profiling (Tompa *et al*., 2002).

### *I.4.4. Epigenetic marks dynamics*

There seems to be interdependence between different epigenetic marks. Evidence from plants suggests the existence of feedback mechanisms between DNA methylation and H3K9me2 such that one promotes the maintenance of the other. DNA methylation at CpG sequences directs H3K9me2 deposition at the heterochromatic chromocentres of somatic nuclei of Arabidopsis (Soppe *et al*., 2002; Tariq *et al*., 2003) and in turn this histone methylation directs DNA methylation at GpNpC sequences

(Jackson *et al*., 2002; Malagnac *et al*., 2002). However, in Arabidopsis it was established that these epigenetic marks are not essential for constitutive heterochromatin formation (Jasencakova *et al*., 2003; Tariq *et al*., 2003). Interactions between histone hypoacetylation, H3K9me2 and cytosine methylation have also been point out for spreading heterochromatin over large domains (Richards and Elgin, 2002). In general, HDACs cooperate with HKMTs and DNA methyltransferases (Pikaard and Lawrence, 2002) in order to implement and maintain a chromatin repressed state.

In organisms like *Saccharomyces cerevisiae, Schizosaccharomyces pombe and Caernorhaditis elegans* there is a lack of DNA methylation. For instance, *S. pombe* needs RNAi machinery to induce H3K9 methylation and histone deacetylation and the subsequent recruitment of chromodomain proteins such as heterochromatin protein HP1 for the formation of heterochromatin (Grewal and Jia, 2007). But also in plants in which DNA is methylable there is an involvement of RNAi in establishing the heterochromatic state (Lippman and Martienssen, 2004). Tariq and Paszkowski (2004) proposed a model combining chromatin remodeling, histone methylation, DNA methylation and RNA interference. Briefly, RNAi-directed DNA methylation results in hypermethylation mainly in CG sequences and the nucleosomes with CG methylated DNA and hypoacetylated histones are modified by H3K9 specific HKMT. DNA is then hypermethylated at CNG and CNN sites and the outcome is gene silencing. The step back towards transcriptional activity starts with DNA demethylation by DME and ROS1, and then there is the replacement of H3K9me2 by H3K4me2

### *I.4.4.3. Epigenenetic patterns during development*

In contrast to animals, development in plants is rather plastic and considerably more affected by environmental effects. Besides having a role in cell cycle regulation, a number of studies have shown that plants are capable of adapting their growth and development to environmental factors such as light, temperature, biotic and abiotic stresses through modulation of histone acetylation (reviewed in Chen and Tian, 2007). It seems that the majority of "epigenetic" marks are reversible (Berger, 2007) in response to stress, development and cell differentiation and therefore different tissues can reveal different patterns of epigenetic marks that can also be heritabale stable throughout multiple cell cycles. Unlike mammals, which erase and reset genomic methylation patterns early in embryogenesis (reviewed in Yoder *et al*., 1997), plants can

inherit epigenetic changes through meiosis. Arabidopsis plants with decreased levels of methylation displayed a collection of phenotypic and developmental abnormalities (Kakutani *et al.*, 1995; Finnegan *et al.*, 1996; Ronemus *et al.*, 1996). The dramatic effect of decreasing DNA methylation in vertebrates and plants suggests that it plays an integral role in regulating gene expression during development, either directly or by maintaining the appropriate chromatin structure (reviewed in Finnegan *et al.*, 1998, 2000).

However, a model is emerging where demethylation by DME is required for the expression of imprinted genes in the endosperm. This model is supported by evidences that show that seeds that inherit a mutant maternal *dme* allele no longer show parent-of-origin specific methylation differences on endosperm alleles: the maternal *Medea* allele is as highly methylated as the paternal allele (Gehring *et al*., 2006). In this model, the *dme* gene is expressed in the central cell of the female gametophyte and DME removes 5-mC from maternal alleles of imprinted genes. The paternal alleles are methylated in the pollen grain and remain in that state. After fertilization, differential expression of maternal and paternal alleles is then observed in the endosperm (reviewed in Gehring and Henikof, 2007). Also the downregulation of the *Flowering Locus C* genes needed for the transition of the apical meristem to a reproductive fate in Arabidopsis (Sheldon *et al*., 1999) can be achieved by vernalization and also by hypomethylation, although through different mechanisms (Finnegan *et al*., 2005).

Although the establishment of modifications of histones is important for development and cell differentiation, it is now clear that the removal of these modifications is also important. The analysis of multiple organisms, including corn, fungus, yeast, *C. elegans, Drosophila,* zebrafish and mice demonstrated that histone H3K4 demethylases are essential in development and differentiation (reviewed in Benevolenskaya, 2007).

### *I.4.4.4.Epigenetic marks during cell differentiation: bicellular mature pollen*

A developmental stage where chromatin modifications are strongly accentuated is in cell differentiation of generative and vegetative cells of mature pollen grains.

While in most animals, the germ-line cells diverge from somatic cells during early embryo development and remain as a distinct cell population throughout the life of the animal, in plants life cycle there are distinct vegetative and reproductive phases. Also in higher plants meiotic products do not differentiate directly into gametes instead they undergo mitotic divisions to form the gametes. The first mitosis results in two unequal products, the generative and the vegetative cell, and the second mitosis is an equal division involving only the smaller generative cell, resulting in the formation of the two sperm cells or male gametes. Angiosperm mature pollen grains can be of the bicellular type, more common (~70% of plant families) like in the Fagaceae family, where the second division occurs only within a germinated pollen tube after pollination; or of the tricellular type where the second mitotic division occurs during pollen maturation in the anther (reviewed in McCormick, 1993; Raven *et al*., 1999). Besides having different sizes and composition, the generative and vegetative cells have markedly distinct nuclear configurations. Generally, at maturation final step the generative nucleus acquires an elongated shape together with a condensed chromatin state, while the vegetative nucleus is nearly spheric and presents diffuse chromatin (Tanaka, 1997).

The distribution of cytosine methylation and of some canonical histone modifications have been evaluated in both nuclei of bicellular pollen of herbaceous species (Oakeley *et al*., 1997; Janousek *et al*., 2000; Okada *et al*., 2006). The accentuated chromatin remodelling in both generative and vegetative nuclei has been attributed to the presence of specific histone variants in the generative nucleus (Ueda *et al*., 2000). These authors proposed that these variants could play a role homologous to the protamine of mammalian sperm in chromatin remodelling of male gametes (reviewed in Wouters-Tyrou *et al*., 1998; Lewis *et al*., 2003). In addition, immunofluorescence staining with histone H1-specific antisera revealed a remarkable decrease in the level of histone H1 in vegetative nuclei of angiosperm pollen being its

more relaxed chromatin structure associated with a low level of histone H1 (Tanaka *et al*., 1998).

Differences can also be detected between the levels of cytosine methylation in generative and vegetative nuclei of mature pollen (Oakeley *et al*., 1997; Janousek *et al*., 2000; Janousek *et al*., 2002) as well as in acetylation of histone H4 isoforms (Janousek *et al*., 2000) and dimethylation at lysines 4 and 9 of histones H3 (Okada *et al*., 2006). These epigenetic marks were not correlated to the distintic nuclear morphological changes but to their different potential of transcription regarding cytosine methylation and dimethylation at lysines 4 and 9 of histones H3 (Oakeley et *al*., 1997; Janousek *et al*., 2000; Janousek *et al*., 2002; Okada *et al*., 2006) and potential for replication (Janousek *et al*., 2000).

## *I.5. Nucleolar organizer regions as functional and dynamic domains*

### *I.5.1. The ribosomal DNA that constitutes Nucleolar organizer regions*

The ribosomal DNA (rDNA) encodes for the four RNA components of the ribosomes, named according to their Svedberg coefficient: 5S, 18S, 5.8S, and 25S (in plants) /28S (in mammals). The rDNA is divided in two multigene rRNA families: the 5S rRNA genes which are transcribed outside the nucleolus by RNA Pol III and their transcripts are incorporated in the large subunit of the ribosome in the nucleolus (Highett *et al.*, 1993); and the 18S-5.8S-25S multigene family which is transcribed by RNA Pol I as a 45S primary transcript which is then processed by splicing in the three rRNA components (reviewed in Pikaard, 2000). The transcription of 45S rDNA gives rise to the most distinctive feature of the interphase nucleus, the nucleolus, and for that reason chromosome sites containing these genes are denominated as nucleolar organizer regions-*nor*s (McClintock, 1934 review in Pikaard, 2000). These multimegabase *loci* consists on arrays of tandemly arranged repeating units of 45S rRNA genes, clustered in hundreds or thousands of copies in one or more sites in the genome, at intersticial or

terminal positions in the chromosomes. Each repeat unit is arranged head-to tail, and consists on the coding portion of the ribosomal rDNA cistron, composed by the 18S, 5.8S, and 25S rRNA genes and the internal transcribed (ITS1 and ITS2) sequences, plus the intergenic (IGS) spacer, further divided in a non-transcribed fraction with subrepeats which lengths are often characteristic of a species, and an external transcribed fraction closer to the 18S rRNA gene where lies the gene promoter (Fig. I.2). Each unit varies around 6-13 kb in angiosperms (Zimmer *et al*., 1988; Bellarosa *et al*., 1990) and around 27-40 kb in gymnosperms (Bobola *et al*., 1992; Karvonen *et al*., 1993). The variations among and within species and individuals are due to differences of IGS length, more precisely to the number and dimensions of subrepeats that are found in the IGS (Moss and Stefanovsky, 1995). The IGS is a complex regulatory unit since rDNA transcription regulatory sequences and pre-rRNA processing signals are located in it. In animals and yeast the knowledge of regulatory sequences and binding proteins is much further advanced than in plants. In higher vertebrates for instance, there are important sequence elements that regulate pre-rRNA transcription, such as the rDNA promoter, enhancers, spacer promoters, origin of replication, transcription terminators, and a replication fork barrier that prevents replication forks from colliding with transcribing RNA polymerase I during S phase. (review in Grummt, 2003; McStay, 2006).

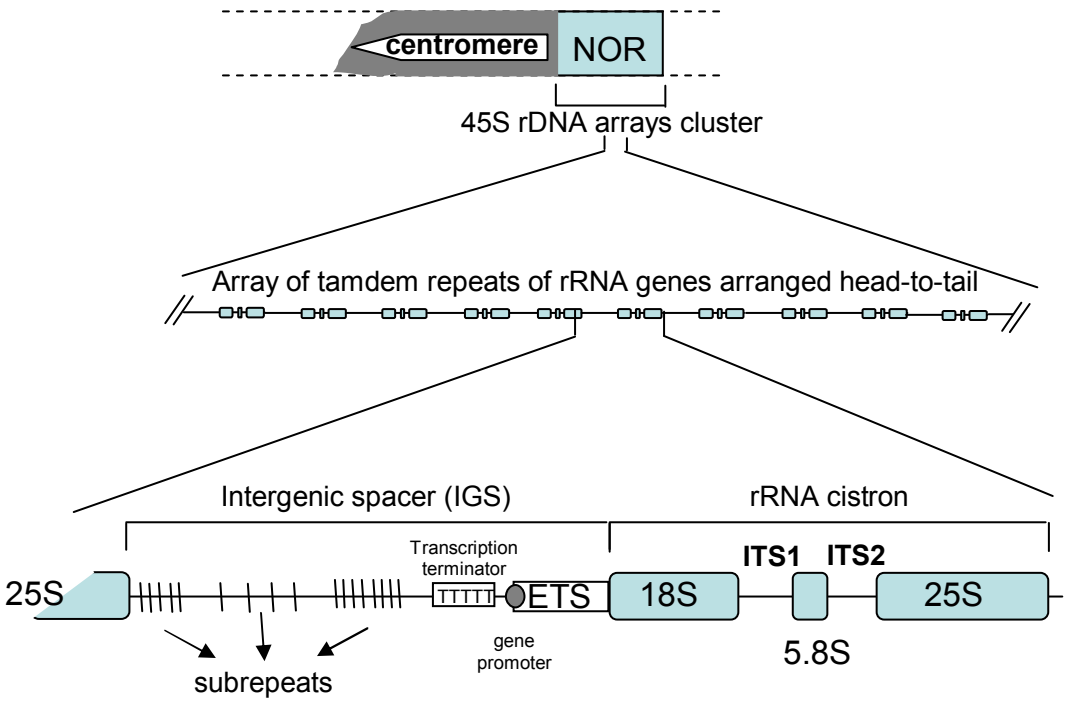



**Figure I.2** – Multilevel organization of 45S rDNA units at NORs (Adapted from Neves *et al*., 2005a)

### I.5.2. NOR chromatin organization

In most eukaryotes the number of rRNA genes is largely redundant in relation to what is required to sustain ribosome assemblage. This is particularly relevant in plants where thousands of copies are present (Ingle *et al*., 1975). In pea, only about 5% of the units are transcribed (González-Melendi *et al*., 2001), suggesting that the majority of the rDNA units remain transcriptionally silent. The first indication that heterochromatic domains are frequently present in nucleolus-organizing regions came from Barbara McClintock (1934) who has proven that a deeply staining body in chromosome 6 of *Zea mays* is associated with the orderly organization of nucleoli (reviewed in Redi *et al*., 2001). A recent model of nucleolar chromatin organization proposes that a subset of rRNA genes is packaged into heterochromatin, which is therefore inaccessible to the transcription machinery, and another fraction is euchromatic and transcribed (Carmo-Fonseca *et al*., 2000). In interphase nuclei, the rDNA units that are transcriptioally silent are detected in many plant species with large and small genomes like cereals and Arabidopsis, respectively, as a chromatin block of condensed rDNA at the nucleolar periphery. In some species such as wheat, condensed *foci* of transcriptional silent rDNA can also appear inside the nucleolus, but the decondensed string that can be seen incorporated in the nucleolus corresponds to the rDNA transcriptional active fraction (reviewed in Neves *et al*., 2005a).

In metaphase rye chromosomes, the active transcriptional rDNA is located at the secondary constrictions of nucleolar chromosomes (Caperta *et al*., 2007), while the untranscribed rDNA units reside as a condensed block at the centromere proximal NOR domain (Caperta *et al*., 2002). Also in *Saccharomyces cerevisiae* it was demonstrated that rDNA silencing is spread only into the centromere proximal sequences (Buck *et al*., 2002). In other species like in wheat, there are also nucleolar chromosomes with two condensed blocks (centromere-proximal and distal blocks) (Morais-Cecílio *et al*., 2000).

A model of NOR organization comprising interphase and metaphase with three ribosomal chromatin states was proposed: the heterochromatic state and two open euchromatic states, which are either transcribed in interphase or nontranscribing in the secondary constriction of the metaphase chromosomes (Huang *et al*., 2006). On the other hand, Neves *et al*. (2005a) emphasizes the heterochromatic nature of 45S rDNA, being its most evident feature their repetitive nature, and how the condensed (inactive)

or decondensed (active) organization can be modulated through genomic environment and developmental controls, through epigenetic mechanisms.

### *I.5.3. NORs expression patterns*

In cycling cells 45S rDNA is expressed from telophase to the end of G2 phase and repressed during mitosis (Prescott, 1964). However not all NORs *loci* are active in the same genome, and within a *locus* only the rRNA genes that are co-localized with the transcription machinery are active (Roussel *et al*., 1996). NORs that were active in nucleolus formation retain in the subsequent metaphase some of the proteins related with transcripton allowing their reactivation at telophase. Although some metaphasic NORs present a decondensed chromatin state or secondary constriction they are transcriptional silent due to post-translational modifications of the transcription machinery (Chen *et al*., 2005). The metaphasic NORs that were transcribed during the preceding interphase as well as nucleoli in interphase nuclei can be detected with silver (Ag-NOR) staining techniques (Goodpasture and Bloom, 1975). Silver staining distinctly labels argyrophilic proteins like RNA polymerase I transcription machinery, including nucleolin, the RNA polymerase I large subunit RPA195, and the vertebrate transcription factor UBF. This label is due to the role of these proteins in reducing the silver solution to metallic silver under acid conditions (Roussel *et al*., 1996; Dundr *et al*., 1997, 2000; Dundr and Olson, 1998).

Resting and growing cells display differential expression of NORs probably due to the fact that protein synthesis is not occurring at a same rate in these two cell populations (Jacob, 1995; Moss and Stefanovsky, 2002). Pikaard (2002) showed that rRNA transcription in non-growing cells is almost undetectable while in growing cells rDNA *loci* may account for 40%-80% of total transcription. The dimensions of nucleoli also reflect rDNA activity, since quiescent cells or those with nucleolar inactivation following induced transcriptional arrest have small and compact nucleoli, in which the distinct morphological components detected through electron microscopy tend to aggregate into adjacent blocks (Shaw and Jordan, 1995). Differential expression of NORs can also occur in the same cell. Silver staining on rye metaphase chromosome with equivalent number of rDNA copies revealed differential expression between homologous NORs which revealed to be independent of parental origin (Caperta *et al*.,

2002). In a marked way, differential NOR expression also occurs in cells of interspecific plant hybrids, displaying nucleolar dominance, i.e. the preferential transcription of the NORs of one parent with partial or total inactivation of the NORs of the other. A model has been proposed for nucleolar dominance in which there is a preferential inactivation through heterochromatization of NORs of the larger parental genome (Viegas *et al*., 2002).

### *I.5.4. Epigenetic modulation of NORs*

In most species a large fraction of rDNA repeats are maintained in a heterochromatic state since only a subset of rDNA repeats are transcribed. The first hint revealing that rDNA clusters could be epigenetically controlled was achieved by Flavell and co-workers (1988) revealing that active *loci* present a higher proportion of rRNA genes with unmethylated cytosine residues in comparison with inactive *loci*. Studies unravelling the mechanisms beneath the nucleolar dominance phenomenon were performed in F1 wheat X rye hybrids where the transcription of the under-dominant rye NOR was re-established after treatment with 5-azacytidine (Vieira *et al*., 1990), a base analogue that prevents cytosine methylation. Studying other hybrid species it was also possible to reactivate repressed rRNA genes by promoting histone hyperacetylation either by using trichostatin A (TSA), a chemical agent that inhibits histone deacetylation (Chen and Pikaard, 1997a), or using transgene plants RNAi-mediated knockdown of histone deacetylases (HDACs) (Lawrence *et al*., 2004). It was then postulated that cytosine methylation and histone deacetlytion are partners that act in the same repression pathway because 5-azacytidine and TSA together were no more effective than alone in derepressing underdominant rRNA genes (Chen and Pikaard, 1997a, b). Further studies using *Arabidopsis suecica*, a natural allopolyploid that displays nucleolar dominance, were able to associate the different types of interphase rDNA domains with several epigenetic marks. A perfect correlation between inactive NORs and co-localization with H3K9met2 was found. In addition, the simultaneous presence of H3K4met3 and H3K9met2 was observed in active NOR *loci*, revealing the double nature of these NORs where a fraction is transcriptional active (H3K4met3), and the other (H3K9met2) remains highly condensed and transcriptionally inactive (Lawrence *et al*., 2004). The activation mechanism of rRNA genes should include the loss of DNA

methylation and H3K9 methylation as well as the accumulation of histone acetylation, although the sequence of events in plants is still not clear, but findings in Arabidopsis concerning the molecular basis for nucleolar dominance fit well with studies of rRNA gene regulation in other systems, particularly in mammals (reviewed in Preuss and Pikaard, 2007). A key component of the rRNA gene regulation in mammals is the activity of the specialized chromatin remodelling complex known as NoRC (Nucleolar Remodelling Complex) (Strohner *et al*., 2001). This complex mediates rRNA gene silencing by recruiting DNA methyltransferase and histone deacetylase to the rDNA promoter region (Santoro *et al*., 2002). The order of events was well established in this system by Santoro and Grummt (2005) since they demonstrated that NoRC-mediated DNA methylation depends in *a priori* of histone deacetylation as *de novo* DNA methylation did not occur in the presence of TSA. These authors also showed that in this system, histone deacetylation and H3K9met are not sufficient marks for rRNA gene silencing, being the cytosine methylation the final step in the chain of events that silence rRNA gene transcription.

# *II. Materials and Methods*

## II.1. Plant materials

### II.1.1. Roots, leaves and pollen collection

In the present work seeds from several species of three genera of the Fagaceae family were used. The species and their provenance are described in Table II.1.

The seeds of these species are all recalcitrant and cannot tolerate desiccation below minimum moisture content and therefore have to be stored in moisture controlled conditions until germination. Seeds were germinated at 24ºC in moist sand, and seminal roots were excised when they reached about 5 centimetres. Seedlings were maintained in moist sand until the development of new adventitious roots and further potted and maintained in the greenhouse. The cytological analysis of meristematic cells was mainly performed on seminal root tips, but adventitious roots were also collected from pots when seminal roots were no longer viable.

For genomic DNA isolation, young leaves were collected from young seedlings and immediately immersed in liquid nitrogen and stored at -80º until DNA extraction.

Catkins containing flowers after anthesis were collected from *Quercus suber* trees during the flowering season. Pollen was isolated from anthers using a copper sieve and kept over dry silica gel at room temperature until being used.

**Table II.1** – Species used in this work with their provenances

| Genus | Species | Provenance |
|---|---|---|
| *Quercus* Linnaeus subgenus *Quercus* | • Evergreen<br>*Quercus suber* Linnaeus<br>*Quercus ilex* L. subsp *rotundifolia* (Lam.) | Mediterranean<br>Acorns kindly provided by Dr. Hachemi Merouani collected in Alcácer do Sal (Portugal)<br>Pollen collected in Sintra (Portugal) |
| | • Deciduous<br>*Quercus acutissima* Carruthers<br>*Quercus serrata* Murray | Asian<br>Acorns kindly provided by Dr. Atsuchi Sakai (Japan) collected in Japan |
| *Castanea* Miller | • Deciduous<br>*Castanea sativa* Miller | European<br>Chestnuts collected in Trás-os-Montes (Portugal) |
| | *Castanea mollissima* Blume | Asian<br>Chestnuts kindly provided by Prof. Carlos Abreu from the Botanical Garden of Universidade de Trás-os-Montes e Alto Douro |
| | *Castanea crenata* Sieb.et Zucc. | Asian<br>Chestnuts kindly provided by Dr. Atsuchi Sakai (Japan) |
| *Fagus* Linnaeus | • Deciduous<br>*Fagus sylvatica* | European<br>Seeds and seedlings kindly provided by Engª Isabel Silvestre Viveiros (nursery) St. Isidro (Portugal) |

### II.1.2. C-mitotic treatment and fixation of root tips

For mitotic chromosome analysis in both seminal and adventitious roots, c-metaphases were induced by root incubation in α–bromonaphthalene-saturated solution for 3-4 hours at room temperature.

Successful molecular cytological experiments require adequately preserved material:

For fluorescence *in situ* hybridization, after c-mitotic treatment, the roots were washed in water and then fixed in fresh ethanol: glacial acetic acid (3:1 v/v), a protein-precipitating fixative, for at least 10 hrs at room temperature. The fixative solution was then renewed and root tips were stored at -20ºC until being used.

For *in situ* immunolocalization of modified histones, roots were fixed in fresh 4% (w/v) formaldehyde solution, a cross-linking fixative that preserves proteins, prepared from paraformaldehyde in 1 x PBS [Phosphate buffered saline: 4.3 mM $Na_2PO_4$, 137 mM NaCl, 2.7 mM KCl, 1.4 mM $KH_2PO_4$, pH adjusted to 7.4], for 30 min at room temperature and immediately used in the subsequent treatments.

For silver staining, roots previously fixed in ethanol acetic acid were transferred to fresh FAA fixative (50% ethanol - 37% formaldehyde - glacial acetic acid, 18:1:1 v/v/v) and maintained at -20ºC for 2 to 3 days before use.

## II.2. Isolation of genomic DNA sequences

### II.2.1. Amplification of DNA sequences by polymerase chain reaction (PCR)

The PCR technology invented by Mullis and Faloona (1987) is a rapid procedure for *in vitro* enzymatic amplification of a specific DNA sequence using multiple cycles of DNA denaturation, primer annealing and DNA replication in the presence of a thermostable DNA polymerase. This technique has been constantly used throughout this work to amplify DNA sequences from several sources like, AFLP fragments, genomic DNA, plasmid DNA as well as to simultaneously label and amplify sequences

producing probes for fluorescent *in situ* hybridization (FISH). Below is an example of one application of the PCR technique (Table II.2). Primers designed to amplify the intergenic spacer of the 45S rDNA repeat unit were used in genomic DNA from Fagaceae spp.

**Tables II.2**- ArabrDNA (Amplification of the IGS of the 45S rDNA unit):

| **PCR Reaction** | |
|---|---|
| Genomic DNA (100ng/µl) | 1 µl |
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 1.5µl |
| dNTPs (10mM) | 1.5 µl |
| Primer Arab25F (50mM) (5'-TGAGATTCAGCCCTTTGTCGCTAAG-3') | 1 µl |
| Primer Arab18R (50mM) (5'-AGACAAGCATATGACTACTGGCAGG-3') | 1 µl |
| Distilled water | 36.75 µl |
| Taq polymerase 5U | 0.75 µl |

| **PCR Program** |
|---|
| 95ºC for 4 min |
| Perform 30 cycles at: |
| 94ºC for 1 min<br>62ºC for 1 min<br>72ºC for 2,30 min |
| 72ºC for 10 min |
| Hold temperature at 4ºC |

PCR products were always checked for size and concentration through agarose gel electrophoresis and then purified using the High Pure PCR Product Purification kit (Roche) according to manufacturer's instructions to remove unincorporated nucleotides, enzymes and salt. Purified unlabelled PCR products were directly sequenced or cloned prior to sequencing reactions.

### *II.2.2. Amplified fragment length polymorphism (AFLP)*

AFLP is a DNA based marker technique developed by Vos *et al*. (1995). The AFLPs are generated by complete restriction endonuclease digestion of total genomic DNA, followed by selective PCR amplification and electrophoresis of a subset of the fragments, resulting in a unique, reproducible DNA fingerprint (or profile) for each

individual. Although this technique is mainly applied in population genetics, linkage mapping and parentage analyses (Meudt and Clarke, 2007), in this work it was used according Reamon-Buttner *et al*. (1999) for repetitive sequences isolation to be used as probes for FISH analysis.

#### *II.2.2.5. Restriction endonuclease digestion and Ligation of adaptors*

*Quercus ilex* subsp *rotundifolia* (250 ng) was digested for 2h at 37ºC with 2,5 U of *Eco*R I and *Mse* I (Invitrogen), 1X reaction buffer (Invitrogen) in a 25 µl reaction volume adjusted with distilled water. These restriction enzymes generate small DNA fragments with optimal size range (<1 kb) for separation on denaturing polyacrylamide gels. After inactivation of restriction endonucleases by incubation at 70º C for 15 min, 25 µl of ligation mix was added to the same tube. This ligation mix consisted on 24 µl of the adapter ligation solution [*Eco*R I/ *Mse* I adapters, 0.4 mM ATP, 10 mM Tris-HCl(pH7.5), 10 mM Mg-acetate, 50 mM K-acetate] and 1U T4 DNA ligase (Invitrogen). The reaction was incubated at 20ºC for 2h. The ligated DNA was diluted 1:10 in TE buffer [10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA] and stored at -20ºC until preamplification. These adapter sequences flanking variable genomic DNA sequences serve as primer binding sites on these restriction fragments.

#### *II.2.2.6. Preamplification reactions and selective AFLP amplification*

The polymerase chain reaction (PCR), technique described in 2.1, is performed in two consecutive reactions. In the preamplification reaction (Tables II.3) the ligated fragments are amplified with an *Mse* I primer containing one selective nucleotide and an *Eco*R I primer containing no selective nucleotides. The PCR products of the preamplification reaction are diluted 1:50 in TE buffer and used as template for the selective amplification (Tables II.4) using two AFLP primers: *Mse* I primer containing three selective nucleotides and an *Eco*R I primer containing two selective nucleotides.

Two combinations were used: *Mse* I-CAA and *Eco*R I-AA; *Mse* I-CAG and E-AC. The two-step amplification strategy results in cleaner and more reproducible fingerprints.

**Tables II.3-**Preamplification reaction of AFLP

| PCR Reaction | |
|---|---|
| 1:10 ligation mix DNA | 5 µl |
| Pre-amp primer mix (AFLP kit) | 40 µl |
| 10X PCR buffer plus Mg | 5 µl |
| Taq DNA polymerase (5 U) | 1 µl |

| PCR Program |
|---|
| Perform 30 cycles at: |
| 94ºC for 30 s<br>56ºC for 60 s<br>72ºC for 60 s |
| Hold temperature at 4ºC |

**Tables II.4-**Selective AFLP amplification (for each primer combination):

| PCR Reaction | |
|---|---|
| 1:50 preamplification reaction DNA | 5 µl |
| 10X PCR buffer plus Mg | 2 µl |
| Primer *Mse* I (includes dNTPs) | 4.5 µl |
| Primer *Eco*R I | 1 µl |
| Distilled water | 7.3 µl |
| Taq DNA polymerase (5 U) | 0.2 µl |

| PCR Program |
|---|
| Perform 30 cycles at: |
| 94ºC for 30 s<br>65ºC for 30 s<br>72ºC for 60 s |
| Lower the annealing temperature each cycle 0,7ºC during 12 cycles |
| Perform 23 cycles at: |
| 94º for 30 s<br>56º for 30 s<br>72º for 60º s |
| Hold temperature at 4ºC |

### II.2.2.7. Separation of amplified fragments on denaturing polyacrylamide gels

The preparations of the (sequencing) gel apparatus involves the addiction of a uniform layer of bind silane in one of the glasses and repel silane in the other glass. The apparatus is mounted vertically and 6% denaturing polyacrylamide (42g Urea, 15ml Acrylamide 40%, 10 ml 10xTBE [Trisbase 108g, Boric acid 55g, 0.5M EDTA pH 8.0 40 ml, distilled water to 1000 ml] and distilled water to a total volume of 100 ml. For the polymerization of the solution, 500 µl 10% fresh APS (Ammonium Persulphate) and 100 µl TEMED (NNN'N'-tetramethylethylenediamine) are inserted with the aid of a syringe, avoiding bubbles. The sharks-tooth comb is inserted upside-down (with half the depth of the teeth facing up). Moister paper was put over the comb (the aperture) to leave the gel overnight. The bottom tray is filled with 1500 ml of 1x TBE, the comb is removed and the gel is warmed up at 55 W for 20 min. The samples are prepared by adding an equal volume of formamide loading buffer (98% formamide, 10 mM EDTA pH 8.0, xylene cyanol and bromophenol) to PCR reactions and DNA marker. The samples are heated for 3 min at 90ºC and immediately place on ice. Formamide and temperature will denature DNA. 7 µl of each sample were loaded onto the wells of a 0.4 mm thin gel. The heater was set to 48ºC, and the gel ran at constant 40 W and an upper threshold of 1500 V until the slower dye was two-thirds down the length of the gel. The apparatus is dismounted and the glass plates are separated. The gel that is sticking to the outer glass is placed in a tray in a 2 L cold fix/stop solution (10% acetic acid) and is agitated gently for 30 min. Afterwards, the gel is rinsed 3x for 2-3min in distilled water using agitation. The gel is transferred to the staining solution (2 g of $AgNO_3$, 3ml of 37% formamide, distilled water up to 2 L) and is agitated during 20 min, then 2 L of developing solution (0.3% w/v $Na_2CO_3$) is pour into the tray with agitation until bands start to appear. When the developing degree has been reached the reaction is stop with a neutralized solution, 2 L of the fix/stop solution (10% acetic acid). The gel is rinsed 2x for 2-3min with distilled water and is left to dry.

### *II.2.2.8. Preparation for probes and sequencing from AFLP fragments*

Distinct DNA fragments could be clearly detected after silver staining, and several were chosen according to reproducibility from two consecutive lanes, and then according to the silver labelling intensity. The desired fragments were excised from the gel with precision using a sharp scalpel, placed into a 1.5 ml microcentrifuge tube and 50 µl of 1x TE buffer was added. DNA fragments were boiled for 5 min and then centrifuged at maximum velocity during 10 min. The supernatant will serve as template for DNA isolation and amplification (Tables II.5). PCR amplification products were either sequenced or used as FISH probes when labelled nucleotides were incorporated (Tables II.6).

**Tables II.5-** AFLP fragment amplification (for each fragment and resulted primer combination):

| **PCR Reaction** | |
|---|---|
| Supernatant of silver staining band | 5 µl |
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 2.5 µl |
| Primer *Mse* I (includes dNTPs) | 9 µl |
| Primer *Eco*R I | 2 µl |
| Distilled water | 25.6 µl |
| Taq polymerase | 0.4 µl |

| **PCR Program** |
|---|
| Perform 30 cycles at: |
| 95ºC for 1 min<br>56ºC for 1 min<br>70ºC for 1 min |
| Hold temperature at 4ºC |

**Tables II.6-** AFLP fragment amplification and labelling (for each band and resulted primer combination):

| PCR Reaction | |
|---|---|
| Supernatant of silver staining band | 5 µl |
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 2.5 µl |
| Primer *Mse* I (includes dNTPs) | 9 µl |
| Primer *Eco*R I | 2 µl |
| Biotin-16-dUTP (1mM) or Digoxigenin-11-dUTP (1mM) | 1,5 µl |
| Distilled water | 24.1 µl |
| Taq polymerase | 0.4 µl |

| PCR Program |
|---|
| Perform 30 cycles at: |
| 95ºC for 1 min<br>56ºC for 1 min<br>70ºC for 1 min |
| Hold temperature at 4ºC |

### *II.2.3. DNA Cloning*

Purified PCR products were inserted into a plasmid vector and amplified in the host cell of *Escherichia coli* using the TOPO TA cloning kit (Invitrogen) according to the manufacturer's instructions. Briefly, PCR products bearing deoxyadenosines at 3' ends were ligated to the linearized vector pCR2.1 containing 3' deoxythymidine residues by the action of T4 DNA ligase. The pCR2.1 plasmid vector contains among other features, the cloning site at the lacZα gene, ampicillin resistance ORF, pUC origin and M13 priming sites. The construct was transformed into chemically competent *E. coli* strain TOP10F'cells, which express the lac repressor ($lacI^q$), by a heat shock of 30 sec at 42º C. Bacteria were grown in the presence of S.O.C. medium [2% tryptone; 0.5% yeast extract; 10mM NaCl; 2.5 mM KCl; 10mM $MgCl_2$; 10 mM $MgSO_4$; 20 mM glucose] by shaking the vials at 37º C for 1 hour at 225 rpm. Transformants were analysed in sterilised LB plates [L-Broth liquid medium (1% Tryptone, 0.5% Yeast Extract, 1% NaCl, pH adjusted to 7) plus 15 g/L agar] containing 100 µg/ml of

ampicillin, 40 μg/ml of X-Gal and 100 μg/ml of IPTG. White colonies, supposedly carrying bacteria cells with recombinant DNA, were picked, with a sterile toothpick, and a further selection was done on new LB plates to eliminate false positives. At least 10 white transformants were checked to confirm the presence and sizes of the insert by a PCR reaction with M13 primers (Tables II.7).

**Tables II.7-**Confirm cloned insert by M13 primers:

| **PCR Reaction** | |
|---|---|
| Picked bacteria from colony | |
| 10X PCR buffer | 1 μl |
| $MgCl_2$ (50mM) | 0.3 μl |
| dNTPs (25mM) | 0.08 μl |
| Primer M13 cw (100mM) (5'-TGTAAAACGACGGCCAGT-3') | 0.1 μl |
| Primer M13 ccw (100mM) (5'-CAGGAAACAGCTATGACC-3') | 0.1 μl |
| Distilled water | 8.22 μl |
| Taq polymerase | 0.1 μl |

| **PCR Program** |
|---|
| 94ºC for 10 min |
| Perform 30 cycles at: |
| 94ºC for 1 min<br>55ºC for 1 min<br>72ºC for 1 min |
| 72ºC for 1 min |
| Hold temperature at 4ºC |

The single positive colonies were picked and grown overnight in 5 ml sterilised L-Broth liquid medium containing 100 μg/ml of ampicillin. Glycerol stocks were produced with aliquots containing 85% of culture and 15% of sterile glycerol and were stored at -80ºC for culture preservation.

### II.2.4. Isolation and purification of plasmid DNA

Isolation and purification of plasmid DNA from bacteria is a standard procedure that gives enough purified plasmid DNA (minipreps) and was performed with Wizard™ Plus Minipreps System (Promega) according to the manufacturer's instructions.

Glycerol stocks containing transformed bacteria cells with a given plasmid DNA were used for minipreparation of plasmid DNA. 30µl of glycerol stock was used to inoculate sterilised 10 ml L-Broth liquid medium. 12µl of the ampicillin (50mg/ml) was added to each tube. The tubes were incubated overnight at 37ºC on a shaking platform (200 cycles/min) to allow cell growth. Since the plasmid with the desired DNA also contains the ampicillin marker gene only bacteria carrying the plasmid will survive and multiply. Ten ml of the overnight culture were divided into 1.5 ml microcentrifuge tubes and centrifuged at 3000 rpm for 15 min at room temperature. Supernatants were discarded and the pellets were drained to remove the medium. Pellets were re-suspended in 200 µl of re-suspension solution (50 mM Tris.HCl pH 7.5, 10 mM EDTA, 100 µg/ml RNase A). 200 µl of lysis solution (0.2 M NaOH, 1% SDS) was added and mixed by inversion until the cell suspension cleared, followed by the addition of 200 µl of neutralization solution (2.55 M KOAc pH 4.8) and mixed again by inversion. The tubes were centrifuged at 12 000 g for 5 min and the supernatants (containing the DNA) were collected in clean tubes. 1 ml of the Wizard™ DNA resin was added to each tube, and a Wizard™ mini-column was prepared: the plunger from a disposable syringe was removed and a Wizard™ column was attached to the syringe nozzle. The resin/DNA mixture was pipetted into the barrel, the syringe plunger was reinserted and the mixture was gently pushed through the mini-column. The mini-column was washed twice with 2 ml of the column washing solution (200 mM NaCl, 20mM Tris.HCl pH 7.5, 5mM EDTA diluted with 95 % ethanol 1:1). To dry the resins containing the DNA the column was removed from the syringe, transferred to a clean tube and centrifuged for 20 sec at 12000 g. The mini-column was then transferred to a clean eppendorf, and 50 µl of 1x TE buffer (prepared by diluting 1: 10 the stock solution 10 x TE buffer: 100 mM Trizma Base, 10 mM EDTA.$2H_2O$, adjust pH to 8.0) was applied to the column and left for 1 min at room temperature to elute the DNA from the resin. The mini-column was centrifuged, as for the washing step, and the buffer containing the DNA was recovered. The plasmid DNA solutions were stored at -20 ºC.

### II.2.5. Bioinformatics Sequence analysis

After sequencing, the data obtained was analysed trough bioinformatics tools. Bioinformatics sequence analysis was based on Genbank available tools BLASTn, BLASTx (Altschul *et al*., 1997). The program Primer premier V was used to design primers. AT and GC content was analysed with Res Stat 1.0: Count Nucleotide Frequencies tool.

## II.3. Cytological preparations

The well spread and preserved chromosomes and nuclei allowed cytogenetic analysis including chromosome counting and length measurements as well as mapping of specific repetitive DNA sequences by fluorescent *in situ* hybridization, evaluation of cytosine methylation and histone modifications patterns by fluorescent immunodetection and active 45 S rDNA *loci* detection by silver staining.

### II.3.1. Pretreatment of slides

Glass microscope slides were previously washed in 3% (v/v) Decon detergent for 30 min, and rinsed thoroughly with distilled water. To improve adhesion of the small nuclei and chromosomes to the glass surface and prevent material loss in the subsequent treatments, slides were coated with a freshly prepared solution of 2% (v/v) APTES (3-aminopropyltriethoxy-silane, Sigma) in acetone for 10 sec, followed by a brief wash in acetone and further wash in water, and air-dried.

### II.3.2. Meristematic nuclei and chromosomes preparations by drop technique

The drop technique was chosen for cytological preparations of meristematic root-tip cells due to the small size of chromosomes and dense cytoplasm of cells. In this

method, fixed material is digested with enzymes and then resuspended in alcohol: acetic acid for dropping. The method used here was adapted from the described in Zoldoš *et al*. (1999). Root tips fixed in ethanol-acetic acid were first washed in the enzyme buffer [0.03 % EDTA (disodium ethylene diamine tetracetate) in 2 x SSC (Saline Sodium Citrate: 0.3 M NaCl, 0.03 M sodium citrate in distilled water, pH adjusted to 7), pH adjusted to 4.2] for 20 min to remove fixative. Non-meristematic tissue was removed with a razor blade under a stereo microscope and only the terminal portion (ca.3 mm) without root cap was transfer to the enzymatic mixture for partial digestion of the cell walls [Enzymatic mixture: 2% (w/v) cellulase 'Onozuka' R10 (from *Trichoderma viride*, Serva), 3% (v/v) pectinase (from *Aspergillus niger*, solution in glycerol, Sigma), 0.3% pectolyase Y-23 (from *Aspergillus japonicus,* Sigma), 0.03% EDTA in 2 x SSC, pH 4.2] and incubated at 37°C for 3 h. The material was completely dispersed with a micropipet and further incubated for an additional 30 min at 37ºC. The lysate of one seminal root tip or three adventitious root tips was then centrifuged at 800 *g* for 5 min, the supernatant was removed and the cellular mass was washed twice in the enzyme buffer, then twice in fresh ice-cold ethanol: glacial acetic acid (3:1 v/v) fixative by centrifugation. Finally, the pellet was resuspended in an appropriate volume of fixative, and about 20 µl was dropped onto an APTES-coated slide and was blown gently, which increases the spreading forces of the fixative giving rise to well-separated chromosomes in complete c-metaphases. Air-dried slides were screened under phase contrast microscopy to check cell density and cytoplasm presence. When cytoplasm was abundant, the fixed suspension was further centrifuged and the fixative replaced by fresh 60% glacial acetic acid during 1 min for cytoplasm removal. After centrifugation, the pellet was resuspended again in an appropriate volume of fixative and the aliquot dropping procedure was carried on.

### *II.3.3. Meristematic nuclei squash preparation*

Squash preparations were used for the distribution analysis of modified histones by fluorescence immunodetection according to the method described by Houben *et al*. (2003) with slight modifications. Root tips fixed in 4% (w/v) formaldehyde solution, after a 10 min wash in enzyme buffer, were subject to enzymatic digestion at 37ºC during 3 h in the same enzymatic mixture described in II.**3.2**. After digestion the

material was washed for 15 min in enzyme buffer. The root tip was transferred to a drop of 1 x PBS placed on a clean slide and there the material was teased out in small fragments with forceps and a needle under a stereo microscope, in order to isolate root tip meristematic cells, discarding all other tissues from the slide. Aliquots of dispersed material fragments were transferred to APTES coated slides and a drop of 45 % acetic acid was added. A cover slip was applied and tapped with a needle to promote cell separation and cytoplasm disruption. The slide was checked under a phase contrast microscope and cells were attached to slides by the dry-ice technique, removing the coverslip with a razor blade.

### *II.3.4. Pollen nuclei preparations*

Nuclei isolation of mature pollen was made according to Pan *et al*. (2004) with modifications. Isolated pollen grains were resuspended in the nuclear isolation buffer for woody plants (WPB) (Loureiro *et al*., 2007), consisting on 0.2 M Tris.HCl, 4 mM MgCl2.6H2O, 2 mM EDTA Na2.2H2O, 86 mM NaCl, 10 mM sodium metabisulfite, 1 % PVP-10, 1 % (v/v) Triton X-100, at pH 7.5. The suspension of pollen grains (0.5 ml of WPB to each 0.01 $cm^3$ of pollen grains) was maintained in ice and was sonicated with an ultrasonic device (Ultrasonic Processor, model UP 50 H, Dr. Hielsher GmbH) at 40 W, using 0.8 cycle with 80% amplitude, for 5 min. The treated solution was filtered through miracloth (Calbiochem) and centrifuged at 500 *g* for 5 min. According to the amount of pellet obtained, some of the supernatant volume was carefully discarded and the pellet resuspended in the remaining volume. Nuclei suspension aliquots were placed on APTES-coated slides and the nuclei were attached to slides by the dry-ice technique, removing the coverslip with a razor blade. Pollen nuclei preparations were then ready for use in immunodetection of modified histones.

For cytosine methylation immunodetection, the nuclei suspension was mixed with an equal volume of the fixative 3:1 (v/v) ethanol: glacial acetic acid mixture, placed on ice during 10 min and centrifuged at 500 *g* during 5 min. The supernatant was discarded and replaced with new 3:1 (v/v) ethanol: glacial acetic acid mixture and the procedure of fixation and centrifugation were repeated. According to the amount of pellet obtained, a volume (approximately 100 µl) of new 3:1 fixative was adjusted and

the pellet resuspended. Aliquots were dropped on to APTES-coated slides and allowed to dry.

For fluorescent *in situ* hybridization (FISH) the procedure was equivalent as for cytosine methylation immunodetection, only prior for the FISH technique, slides were incubated with 200 µl of fresh 4% (w/v) formaldehyde fixative solution during 15 minutes to ensure preservation of the material in the following steps of the technique.

## II.4. Silver staining

Silver staining of chromosome spreads is a cytological method commonly used to detect nucleolar organising regions (NORs) which were transcriptionally active in the previous interphase (Hubbel, 1985; Zurita *et al.*, 1998; Morais-Cecílio *et al.*, 2000; Caperta *et al.*, 2002). The silver staining technique used in the present work was adapted from the salt nylon technique of Stack *et al.* (1991) and was also used for nucleoli counting. Silver staining was performed on nuclei and chromosome preparations from root tips fixed in FAA. Prior to staining slides were incubated at 60 ºC for 3 hrs. Slides were incubated during 8 min in 2x SSC at 60 ºC and then washed several times in distilled water and air dried. 100 % (w/v) $AgNO_3$ solution (w/v) made in distilled water was applied to each slide covering the preparation area. Nylon coverslips (mesh pore diameter ~250 µm) were used and the slides were incubated at 60 ºC for 10-15 min in humid chamber. The reaction was stopped by thoroughly washing the slides in distilled water. Slides were then mounted in immersion oil.

## II.5. Fluorescent *in situ* hybridization technique (FISH)

The development of the DNA *in situ* hybridization technique (Gall and Pardue, 1969; John *et al.*, 1969) marked the transition from the classical cytogenetics era to the modern molecular cytogenetics era. This technique combines the molecular information

of the DNA sequences with their physical analysis. The basic procedure of DNA *in situ* hybridization is the labelling of a DNA probe and its hybridization to cytological preparations. Radiation-based methods were used in probe labelling and signal detection in early techniques. However, such methods were soon replaced by fluorescence-based techniques. The fluorescent DNA: DNA *in situ* hybridization technique (FISH) allows the visualization of specific *loci* throughout the cell cycle and will be used in this work to determine abundance and distribution of repetitive sequences making up the majority of most genomes. The analysis of the FISH patterns will provide information for investigations on genome organization, karyotyping, phylogenetic analysis and ploidy assessment in Fagaceae species. The FISH method was carried out mostly according to the procedures described in Schwarzacher and Heslop-Harrison (2000).

### *II.5.1. DNA probes*

Cloned DNA sequences and polymerase chain reaction (PCR) products were used as probes in this work. The list of DNA probes is described below:

- pTa71, containing a 9 kb EcoRI fragment from the highly repeated 45S rDNA sequence isolated from wheat (Gerlach and Bedbrook, 1979), cloned into the plasmid pUC19.
- pTa794, containing a 410 bp fragment of the repeated 5S rDNA sequence isolated from wheat (Gerlach and Dyer, 1980) and cloned into the plasmid pBR322.
- 5SQsu, a 353 bp fragment (including 89 bp of the conserved gene sequence and 264 bp of the non-transcribed spacer) of the repeated 5S rDNA sequence obtained by polymerase chain reaction from genomic DNA of *Quercus suber*.
- telomeric repeats $(TTTAGGG)_n$
- AFLP probes (Rot2; Rot8; Rot10; Rot20; Rot28)
- Intergenic rDNA sequences isolated from *Castanea sativa* (476 bp, 351bp and ~300 bp) and *Fagus sylvatica* (305 bp).

### II.5.2. DNA probe labelling by nick translation or polymerase chain reaction (PCR) techniques

Detection of DNA hybrids molecules employs the use of modified (labelled) nucleotides. DNA was labelled by the primary incorporation of non-fluorescent nucleotides, either biotin-11-dUTP (Sigma) or digoxigenin-11-dUTP (Boehringer Mannheim). Two methods were used for probe labelling. Nick translation method is used to label large cloned inserts and thus pTa71 was the only probe labelled by this method. Polymerase chain reaction (PCR) has become a widely used method to label as it amplifies a DNA sequence smaller than 1 kb and therefore all other probes were labelled by this method. The nick translation technique is based on the activity of two enzymes: DNase I, an endonuclease that hydrolyzes each strand of a double stranded DNA molecule at various sites (nicks); and *E. coli* DNA polymerase I which adds nucleotides (labelled or unlabelled) to the 3'-end of the nick copying the template in the opposite DNA strand. The labelled fragments obtained by this method show a length distribution in the range of 200 to 500 nucleotides, which is the optimal range for *in situ* hybridization experiments. The nick translation kit (Roche) was used following the standard protocol; with the exception of labelled dNTPs all reagents are supplied. In short, to an ice cold microcentrifuge tube were added 1 μg of plasmid DNA, 5 μl of dNTP mixture [0.08 mM of labelled dNTP (digoxigenin-11-dUTP or biotin-11-dUTP), 0.17 mM dTTP, 0.25 mM dATP, 0.25 mM dCTP, 0.25 mM dGTP], sterile water up to 16 μl and finally 4 μl of Nick Translation Mix containing 5 x conc stabilized reaction buffer in 50% glycerol (v/v) and DNA-polymerase I / DNase I enzyme mixture. The reaction which occurred at 15ºC for 1h was stopped by adding 1 μl 0.5 M EDTA (pH 8.0) and heating to 65ºC for 10 min.

The incorporation of labelled nucleotides into the probe was subsequently checked by a test-blot analysis adapted from a protocol presented by Schwarzacher and Heslop-Harrison (2000). A piece of Hybond N+ membrane (Amersham) was soaked in buffer no.1 (0.1 M Tris.HCl pH 7.5, 0.15 M NaCl) and dried between filter papers. One μl probe sample was loaded in the membrane and allowed to dry for 10 min, the membrane was then transferred to buffer no.1 for 1 min, and incubated in buffer no.2 [0.5% (w/v) blocking agent (Boehringer Mannheim) in buffer no.1] for 30 min at room temperature, shaking gently. The membrane was then incubated at 37ºC for 30 min in a

solution of antibodies for the labels. This solution was prepared using a 1:100 dilution of anti-digoxigenin conjugated to fluorochrome fluorescein isothiocyanate (anti-dig-FITC, Boeringer Mannheim) and a 1:1000 dilution of streptavidin conjugated to fluorochrome cyanine 3 (streptavidin -Cy3, Sigma) in buffer no.1. The membrane was washed in buffer no.1 for 15 min and washed in water and air-dried in the dark. The membrane was placed over a slide and the positively labelled DNA is seen as fluorescent (green or red-according to the conjugated fluorochromes used) dots on the membrane under the epifluorescence microscope with the appropriated filters.

For the PCR labelling of probes, the reaction with respective primers and the inclusion of labelled nucleotides as well as the appropriate thermocycler programs are described for each probe:

**Tables II.8**-Components, concentrations and PCR program used to simultaneously amplify and label the pTa794 plasmid

| **PCR Reaction** | |
|---|---|
| Miniprep (10ng/µl) | 1 µl |
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 1.5 µl |
| dNTPs (10mM) | 1 µl |
| Biotin-16-dUTP or Digoxigenin-11-dUTP (1mM) | 1.5 µl |
| Primer pBR322 *Bam*HI cw (30mM) (5'-CACTATCGACTACGCGATCATGG-3') | 1 µl |
| Primer pBR322 *Bam*HI ccw (30mM) (5'-ATCGGTGATGTCGGCGATATAGG-3') | 1 µl |
| Distilled water | 37.6 µl |
| Taq polymerase | 0.4 µl |

| **PCR Program** |
|---|
| 94ºC for 4 min |
| Perform 30 cycles at: |
| 94ºC for 45 s<br>60ºC for 1 min<br>72ºC for 1 min |
| 72ºC for 3 min |
| Hold temperature at 4ºC |

**Tables II.9**-Components, concentrations and PCR program used to simultaneously amplify and label the Qsu5S sequence from *Q. suber* genomic DNA

| PCR Reaction | |
|---|---|
| Genomic DNA from *Q. suber* (100ng/µl) | 1 µl |
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 1.5 µl |
| dNTPs (10mM) | 1 µl |
| Biotin-16-dUTP or Digoxigenin-11-dUTP (1mM) | 1.5 µl |
| Primer 5SQsu cw (25mM) (5'-ATCCCATCAGAACTCCG-3') | 1 µl |
| Primer 5SQsu ccw (25mM) (5'-GCAACGATGCTCCTTAA-3') | 1 µl |
| Distilled water | 37.6 µl |
| Taq polymerase | 0.4 µl |

| PCR Program |
|---|
| Perform 30 cycles at: |
| 94ºC for 1 min<br>45ºC for 45 s<br>72ºC for 1,30 min |
| 72ºC for 6 min |
| Hold temperature at 4ºC |

**Tables II.10**-Components, concentrations and PCR program used to simultaneously amplify and label telomeric sequences

| PCR Reaction | |
|---|---|
| 10X PCR buffer | 5 µl |
| $MgCl_2$ (50mM) | 1.5 µl |
| dNTPs (2,5mM) | 2 µl |
| Biotin-16-dUTP or Digoxigenin-11-dUTP (1mM) | 1.5 µl |
| Primer TEL cw (50µM) $(5'\text{-TTTAGGG-}3')_5$ | 2 µl |
| Primer TEL ccw (50µM) $(5'\text{-CCCTAAA-}3')_5$ | 2 µl |
| Distilled water | 35.6 µl |
| Taq polymerase | 0.4 µl |

| PCR Program |
|---|
| Perform 10 cycles at: |
| 94ºC for 1 min<br>55ºC for 1 min<br>72ºC for 1 min |
| Perform 30 cycles at: |
| 94ºC for 1 min<br>60ºC for 1 min<br>72ºC for 1.30 min |
| Hold temperature at 4ºC |

**Tables II.11-**Components, concentrations and PCR program used to simultaneously amplify and label intergenic rDNA sequences isolated from *Castanea sativa* and *Fagus sylvatica*

| PCR Reaction | |
|---|---|
| Purified DNA from agarose band (50ng/µl) | 1 µl |
| 10X PCR buffer | 2.5 µl |
| $MgCl_2$ (50mM) | 0.75 µl |
| dNTPs (10mM) | 0.5 µl |
| Biotin-16-dUTP or Digoxigenin-11-dUTP (1mM) | 0.75 µl |
| Primer Arab25F (50mM) (5'-TGAGATTCAGCCCTTTGTCGCTAAG-3') | 0.5 µl |
| Primer Arab18R (50mM) (5'-AGACAAGCATATGACTACTGGCAGG-3') | 0.5 µl |
| Distilled water | 21.5 µl |
| Taq polymerase | 0.25 µl |

| PCR Program |
|---|
| Perform 30 cycles at: |
| 94ºC for 1 min<br>45ºC for 45 s<br>72ºC for 1,30 min |
| 72ºC for 6 min |
| Hold temperature at 4ºC |

The PCR product obtained was checked for size, concentration and incorporation of labelling by agarose gel electrophoresis. All the probes that reveal positive labelling were purified using the High Pure PCR Product Purification kit (Roche)

### *II.5.3. Pretreatments of slide preparation for DNA in situ hybridization*

A mild digestion of cytoplasm was carried out through pepsin treatment in drop preparations. Slides were incubated with a pepsin solution [5 µg/ml pepsin (3.200-4.500 units/mg) in 0.01 M HCl] for 10 min at 37 ºC in a humid chamber. The pepsin was removed by washing the slides 2 x 5 min in 2 x SSC [Saline Sodium Citrate: 1:10 dilution of a stock solution 20 x SSC (3 M NaCl, 0.3 M sodium citrate in distilled water, pH adjusted to 7)].

After the pepsin treatment the drop preparations were submitted to RNA digestion to remove RNA that could bind to the probe. The slides were then incubated

with a RNase A solution [100 µg/ml of DNase-free RNase (Sigma) in 10 mM Tris.HCl pH 8.0] at 37 ºC for 1 h in humid chamber, followed by washing in 2 x SSC.

### *II.5.4. In situ hybridization mixture and hybridization conditions*

The hybridization mixture was prepared in order to optimise hybridization conditions and to control the stringency at which the experiment was carried out. Stringency is a parameter that determines the percentage of nucleotides that correctly match in the hybrid DNA molecule. In the present work, a stringency of 76% was used (Schwarzacher and Heslop-Harrison, 2000) with a final concentration of 50% (v/v) formamide and 2x SSC in the hybridization mixture, at 37ºC. Besides the DNA probes, the hybridization mixture also contains sodium dodecyl sulphate (SDS) that contributes to the permeabilization of the material; salmon sperm DNA as a blocking reagent to prevent non-specific hybridization of the probe; and dextran sulphate to increase the hybridization rate. Table II.12 exemplifies the hybridization mixtures used.

**Table II.12** Components and concentrations of a probe hybridization mixture.

| Stock Solution | Final concentration in the hybridization mixture |
|---|---|
| 100 % formamide | 50 % |
| 20 x SSC | 2x |
| 50 % (w/v) dextran sulphate | 10 % |
| 10 % (w/v) SDS | 0.2 % or 0.1 % |
| salmon sperm DNA (5 µg/µl) | 1.5 µg/µl |
| probe(s) | x µl |
| sterile water | Adjust to final volume |

The amount of probe(s) used depends on the probe type. The most effective hybridization experiments were performed using about 2.5 ng/µl for pTa71, 3.5 ng/µl for pTa794, and 6ng/µl ng for the telomeric and *Quercus ilex* L. subsp. *rotundifolia* - isolated repetitive sequences.

The denaturation of the hybridization mixture was carried out at 85 ºC for 7 min in water bath, and then cooled in ice for 5 min and immediately applied to the preparations and covered with a plastic coverslip. The slides were placed in a modified thermocycler (JMResearch CT 100) and DNA denaturation was performed at 85 ºC for 5 min. The temperature was then slowly reduced to 37ºC and slides were transferred to a humid chamber at 37 ºC for overnight hybridization.

Post-hybridization washes were carried out by two 5 min washes using a 20 % (v/v) of formamide in 0.1 x SSC (1:200 dilution of 20 x SSC - 3 M NaCl, 0.3 M Na citrate) at 42 ºC corresponding to a stringency of 86 % followed by two 5 min washes in 2 x SSC at 42ºC, and two more washes at room temperature. Slides were then washed 2 x 5 min in 4 x SSC/0.2% Tween-20.

### *II.5.5. Detection of hybridization sites*

Antibodies specific to the nucleotide modifications and conjugated with fluorochromes were used for the detection of hybridization sites. Before adding the antibody solution, the slides are pre-incubated in the blocking solution of 5% bovine albumine serum (BSA) in 4 x SSC/0.2% Tween-20 to prevent unspecific binding. Probes labelled with digoxigenin were detected with anti-digoxigenin antibody conjugated to FITC (Boeringer Mannheim); biotin-labelled probes were detected with streptavidin-Cy3 (Sigma). Both antibodies were diluted in 5% BSA in 4 x SSC/0.2% Tween-20, 1:50 and 1:300 respectively. Slide incubation with the antibodies was carried out in humid chamber for 1 h at 37 ºC followed by 3x 5 min washes in 4x SSC/0.2 % Tween-20 at room temperature.

DNA was counterstained with a 2 µg/ml DAPI (4',6-diamidino-2-phenylindole) solution for nuclei and chromosomes visualization and slides were mounted in antifade solution (AF1 Citifluor Ltd.).

## *II.6. Fluorescent in situ immunodetection*

*In situ* immunodetection enables the physical location of molecules in cells and tissues using specific antibodies conjugated with a visible marker. The indirect method was used in this present work, with a secondary antibody labelled with a fluorochrome.

Antibodies rose against epitopes consisting on chemical modifications of histones and DNA were used for the physical location of epigenetic marks in chromatin, enabling comparative investigations on epigenome organization.

### II.6.1. Primary antibodies

The following primary antibodies were used:

- Anti-H3K4me3, a rabbit polyclonal antibody raised against the trimethylation of the lysine 4 of histone H3 aminoacids tail (Abcam AB 8580)
- Anti-H3K9me2, a rabbit polyclonal antibody raised against the dimethylation of lysine 9 of histone H3 aminoacids tail (Abcam AB 7312)
- Anti-H3K9Ac, a rabbit polyclonal antibody raised against the acetylation of lysine 9 of histone H3 aminoacids tail (Upstate 06-942)
- Anti-H4K5,K8,K12,K16Ac, a rabbit polyclonal antibody raised against the acetylation of lysines in the positions 4,5,8,12 and 16 of the aminoacids tail of histone H4 (Abcam AB 2380). This antibody recognises all acetylated isoforms of histone H4, but binds best to the tetra-acetylated form.
- Anti-5mC, a mouse monoclonal antibody raised against 5-methylcytosine, the methylation of cytosines at the 5 position of the pyrimidine ring (Abcam AB 10805)

### II.6.2. Immunodetection of histone modifications

Immunodetection of modified histones was carried out following the procedure described by Houben et al. (2003). To avoid non-specific antibody binding, slides were incubated for 30 min in 8% (w/v) BSA, 0.1% (v/v) Triton X-100 in PBS at room temperature. After washing in 1X PBS, the slides were incubated for 12h at 4º C in a humid chamber with the primary antibodies in the following dilutions: 1:200 (anti-H3K9Ac), 1:500 (anti-H3K9me2 and anti-H4K5,K8,K12,K16Ac) and 1:1000 (anti-H3K4me3) in PBS and 1% BSA. After this incubation the slides were washed in 1X PBS and then incubated in Cy3-conjugated anti-rabbit IgG diluted 1:100 in 1X PBS, 1% BSA for 1h at 37º C. After final washes in 1X PBS, the preparations were mounted in antifade containing DAPI as DNA counterstain.

### II.6.3. Immunodetection of methylated cytosines

The slides were blocked for 30 min in 1% (w/v) BSA in 1X PBST (PBS supplemented with 0.5 % (v/v) Tween 20) at room temperature. After washing in 1X PBST, the slides were incubated for 1 h at 37º C in a humid chamber with anti-5-mC diluted 1:200 in 1% (w/v) BSA in 1X PBST. The slides were further washed in 1X PBST and then incubated in Cy3-conjugated anti-mouse IgG diluted 1:100 in 1X PBS, 1% BSA for 1h at 37º C. Finally, the slides were washed in 1X PBST and mounted in antifade containing DAPI as DNA counterstain.

## II.7. Cell analysis and image acquisition

Mature bicellular pollen grains were spread over a slide without a coverslip to prevent differential squashing and comparisons of sizes were made under a bright field light microscope. Silver staining preparations were also analysed using a bright field light microscope (Axioskop 2, Zeiss). Hybridized and immunodetected preparations were analysed with an epifluorescence microscope (Axioskop 2, Zeiss), and images were collected using an AxioCam digital camera (Zeiss) controlled by AxioVision 3.0 and assembled using Adobe Photoshop 6.0 (Adobe Systems Inc.). Measurements were performed using AxioVision measurement module 3.0.0. (Zeiss).

# *III. Intergeneric, inter- and intraspecific variability*

## III.1. Introduction

Divergence and speciation are often accompanied by rearrangements in chromosome complements being the comparative analysis of karyotypes an important tool for evolutionary studies. Each karyotype is defined by the number of chromosomes and their characterization based on their morphology. A graphic presentation is often used to give a better illustration of the chromosomes and their morphological features attributed by statistical parameters of values based on their measurements. Classification of chromosomes by centromere position is a basic feature of karyotype analysis. Karyological classifications vary however from study to study (review Borzan and Schlarbaum, 2004). Levan and coworkers (1964) tried to develop precise standards for the nomenclature of chromosomes, for a better standardization among research community. Once established, karyotypes can be compared between species and one way to perform such comparison in evolutionary terms is through the concept of symmetric *vs*. asymmetric karyotypes. When there is predominance of the centromere position at the median point and region and chromosomes with approximately similar sizes, it is considered a symmetric karyotype. The shift of the centromere position to other categories such as submedian and subterminal region and/or with differences in the relative size between the chromosomes of the complement, thus making the karyotype more heterogeneous, constitutes an asymmetric karyotype. Some karyomorphological parameters can be used to infer asymmetrical indexes (reviewed in Paszko, 2006). Relative lengths of chromosomes can be used to estimate the molecular size of each chromosome and with the help of FISH BAC landmarks size-based karyotypes can be used to order molecular markers, measure physical genome distances and for the integration of linkage groups nomenclature and chromosomes designations, which enables a unique chromosome nomenclature for a given species (Kim *et al*., 2005).

In the karyotype establishment chromosome morphometric measurements are important to discover the structure of chromosomes but are often not sufficient for the unambiguous identification of individual chromosomes and their homologues. Some karyotypes are composed of chromosomes with similar sizes and morphology, and then there is a need for additional chromosomal landmarks to discriminate the homologues

pairs of the complement. Repetitive DNA sequence can generate unique FISH patterns on individual chromosomes for karyotyping and phylogenetic analysis (Jiang and Gill, 2006). From these sequences, ribosomal genes and telomere repeats present in all eukaryotic genomes are very helpful since their number and position may vary between taxa. The genes encoding the 18S-5.8S-25S ribosomal RNA (45S rDNA) and 5S ribosomal RNA (5S rDNA) have become the most widely used FISH markers in taxonomic and phylogenetic studies, due to their highly conserved gene sequence, tandem repeated nature, and large number of copies per *locus*, which makes them easy to detect in a wide range of species.

A molecular-cytogenetic analysis of European and American *Quercus* species revealed a high degree of conservation, concerning chromosome number and rDNA *loci* number and mapping (Zoldoš *et al*., 1999), unravelling a small evolutionary rate within this genus. The Fagaceae family has probably an Asiatic origin, being suggested that the basal genus *Fagus* appeared first in Asia, and therefore the Asian species seems to be ancient than the European ones (Paffetti *et al*., 2007). It was also established by molecular means that species of genus *Castanea* have an Asian origin followed by a westward migration via Europe to North America (Lang *et al*., 2007).

In order to study the genome evolution in the Fagaceae family, cytotaxonomic comparisons of several European and Asiatic species belonging to the most represented European genera *Quercus*, *Castanea*, and *Fagus* (Manos *et al*., 2001; Li *et al*., 2004) were performed through the evaluation of some karyotype features such as chromosome structure, asymmetry index, and number and position of rDNA *loci*.

## III.2. Materials and Methods

### III.2.1. Plant material

Root-tips were collected mainly from germinated acorns or from seedlings in pots from the following species: *Fagus sylvatica* L., *Quercus suber* L., *Quercus ilex* L. subsp *rotundifolia* (Lam) O. Schwarz, *Quercus acutissima* Carruth., *Quercus serrata* Murray, *Castanea sativa* Mill. *Castanea crenata* Sieb. & Zucc, and *Castanea mollissima* Bl.

Mature pollen grains were obtained by sieving dehiscent anthers from catkins collected from *Quercus suber* L. trees.

### III.2.2. Chromosome preparations

C-metaphases were induced by root treatment with a α–bromonaphthalene-saturated solution for 3-4 hours at room temperature. Root tips were fixed in fresh ethanol: glacial acetic acid (3:1 v/v) and well spread chromosomes were obtained by the drop technique described in section **II.3.2**.

### III.2.3. Pollen preparations

Mature bicellular pollen grains were spread over a slide and the evaluation of pollen grain sizes was performed without a coverslip to prevent differential squashing. Pollen nuclei preparations for fluorescent *in situ* hybridization (FISH) were made according to the procedure described in section **II.3.4**.

### III.2.4. DNA:DNA FISH

pTa71 probe corresponding to the 45S rDNA and pTa794 probe, corresponding to the 5S rDNA were used in all species except in *Quercus suber* L. In this species a specific probe- Qsu5S- probe was produced (for more details see section **II.5.1**). Arabidopsis-like telomeric repeats sequence $(TTTAGGG)_n$, produced according the procedure described in **II.5.2**, was used to detect telomeric regions. Probes labelled

with digoxigenin were detected with anti-digoxigenin antibody conjugated to FITC, and biotin-labelled probes were detected with streptavidin conjugated to Cy3.

### III.2.5. Cell analysis and image acquisition

Hybridized slides were analysed with an epifluorescence microscope (Axioskop 2, Zeiss). Mature pollen grains from *Quercus suber* L were screened over a slide with bright light microscope. All images were collected using an AxioCam digital camera (Zeiss) controlled by AxioVision 3.0 and assembled using Adobe Photoshop 6.0.

### III.2.6. Morphometric analysis

Measurements were performed in individualized metaphase straight chromosomes of different individuals of *Fagus sylvatica*, *Quercus suber* and *Castanea sativa* using AxioVision measurement module 3.0.0. (Zeiss). For each chromosome of the complement the total length, as well as short and long chromosome arms lengths were measured identifying the centromere by the primary constriction in DAPI staining. The data was then exported to Excel spreadsheet and pairs of homologues of each complement were identified through the total length and the centromeric index (CI =100 x length of the short arm / total length) as well as mapping rDNA FISH landmarks when available. Mean and standard deviations were calculated for all metaphase pairs of a species. Also the chromosome relative contribution to the karyotype was estimated regarding its relative length (percentage of chromosome length in the total of the karyotype length (genome length)) and relative DNA content.

Based on Levan *et al.* (1964) chromosomes can be considered as: **M** (centromere at median point) when the CI values falls in the range from 49.0 to 50.0; **m** (centromere at median region) between 37.6 and 48.9; **sm** (centromere at submedian region) between 25.1 and 37.5, and **st** (centromere at subterminal region) between 12.5 and 25.0.

Asymmetry index (AI) was calculated according to Paszko (2006) multiplying the coefficient of variation of chromosome length and the coefficient of variation of centromeric index of the complement.

## III.3. Results

### III.3.1. Karyotypes of *Fagus sylvatica*, *Quercus suber* and *Castanea sativa*

The karyotypes variability between species of the three genera *Fagus*, *Quercus* and *Castanea* was assessed comparing karyotypes, using morphometric analysis as well as rDNA FISH markers. Telomeric repeats were also used to detect karyotype variability in this family concerning copy number and location of these sequences. *Fagus sylvatica, Quercus suber* and *Castanea sativa*, have the same chromosome number 2n= 24 (Fig. III.1 a,b,c).

The total length range of chromosomes was measured on four metaphases cells in *Fagus sylvatica,* six in *Quercus suber* and eight in *Castanea sativa*, although in some cases not all chromosomes could get into account because they were too folded. rDNA FISH landmarks were helpful in identifying homologues pairs and they were mapped in all species at the short arms of the chromosomes (Tables III. 1a, 2a and 3b). The relative contribution for the karyotypes of the biggest and the smallest pairs of chromosomes were, respectively: 10.31 % and 6.13 % in *Fagus sylvatica* (Table III.1a); 12.12 % and 5.89 % in *Quercus suber* (Table III.2 a), and 11.49 % and 6.37 % in *Castanea sativa* (Table III.3 a). The biggest chromosome of *Quercus suber* is more than twice the size of the smallest chromosome, while in *Fagus sylvatica* and *Castanea sativa* it is almost twice the size, but in all species there is a gradation of lengths in the remaining chromosomes. The relative lengths of chromosomes were used to estimate the average molecular size of each chromosome, considering equal chromatin compaction along chromosomes. Estimation of chromosomes DNA amount based in proportionality with chromosome length, show the following variations among the three species analysed: *F. sylvatica* with a genome size of 544 Mbp/C (Gallois *et al*., 1999) show a variation from 33.35 to 56.09 Mbp, *Q. suber* with 931 Mbp/C (Zoldoš *et al*., 1998) displays a range from 54.84 to 112.84 Mbp, and *C. sativa* with 956 Mbp/C (Barow and Meister, 2002) presents chromosomes ranging from 60.9 to 109.84 Mbp (Tables III. 1a, 2a, 3a).

Using the chromosome classification proposed by Levan *et al.* (1964) and based on values of centromeric index, karyotypes of *Fagus sylvatica, Quercus suber* and *Castanea sativa* were characterized according to chromosome type (Table III.1b, 2b and

3b). Chromosomes that display CI mean values that overlap two categories (considering the standard deviation) are indicated with labels for both categories. *Fagus sylvatica* can be characterized as having a karyotype composed by 1M (7$^{th}$ pair), 7m, 3(4)sm (3$^{rd}$, 9$^{th}$ and 10$^{th}$ pairs with certainty and 11$^{th}$ pair with ambiguity) and 1st/sm (11$^{th}$ pair) (Table III.1 b), while the karyotype of *Quercus suber* can be characterized as 1M (10$^{th}$ pair), 3m, 7sm and 1 st/sm (the 5$^{th}$ pair which bears the major terminal 45S rDNA *locus* on the short arm) (Table III.2 b), and the karyotype of *Castanea sativa* can be expressed as 1M (7$^{th}$ pair), 9m, 1m/M (8$^{th}$ pair) and 1sm/m (11$^{th}$ pair) (Table III.3 b). The only karyotype with two ambiguous categories is that of *Castanea sativa* which presents the higher range of chromosome length.

Chi square analysis of the three karyotypes regarding the distribution of chromosome morphometric classes (M, m, sm and st) (Table III.4) shows that the three karyotypes have the same distribution ($\chi 2$ =9.54, df=6, p=0.15).

The karyotype asymmetry index is a good expression of the general morphology of plant chromosomes. Using chromosomes length and centromeric index values, karyotypes asymmetry was assessed through the determination of karyotype asymmetric index (AI), as described by Paszko (2006) The AI value represents the heterogeneity of chromosome length and centromeric position in a given karyotype. As the AI index gets higher, so does karyotype asymmetry. As the index gets lower, it indicates greater karyotype symmetry. Thus, *Quercus suber* presents the most asymmetric (AI=3.60) karyotype and *Castanea sativa* the most symmetric one (AI=1.88). *Fagus sylvatica* has a value of AI= 2.14 (Table III.5). Although the karyomorphology seems more divergent between *Quercus suber* and *Castanea sativa* species, than when compared with *F. sylvatica*, studies of rDNA *loci* distribution revealed that *Castanea sativa* and *Quercus suber* share the same number of rDNA *loci* and similar chromosome locations with one pericentromeric 5S rDNA *locus* and two 45SrDNA *loci* (one pericentromeric with a less intense signal -minor NOR- and a terminal one with a stronger signal -major NOR-) on the short arms (Fig. III.1 b,c,d), while *Fagus sylvatica* have two 5S rDNA *loci* at pericentromeric positions and four 45S rDNA *loci* at terminal positions all located on the chromosomes short arms (Fig. III.1 a,d).

The detection of telomeric sequences was performed with the Arabidopsis-type telomere sequences, and all the species analysed present hybridization signal strictly located at the end of chromosome arms

FISH signals have different intensity being more intense in *C. sativa,* fainter in

*Quercus suber* and faintest in *F. sylvatica* (Fig. III.2). In *C. sativa* the maximum number of signals observed was 96 (4x24 chromosomes), being detected in all chromatids, although with different intensities in some chromosome arms (Fig. III.2 a). In *Q. suber* the maximum number observed was 58, as some chromatids or even chromosome arms lack any signal (Fig. III.2 b). In *Fagus sylvatica* the majority of the signals were only seen in one arm of the chromosomes and therefore the maximum number observed was 13 (Fig. III.2 c).

Using these parameters the karyotype of each species, ordering the chromosomes from the bigger to the smaller was composed, and represented graphically by the ideograms of Fig. III.3.

**Figure III.1** – Fluorescent in situ hybridization in c-metaphases of *Fagus sylvatica* L**. a)**, *Quercus suber* L. **b)**, and *Castanea sativa* Mill. **c)**, with the ribosomal DNA probes as FISH markers in helping detecting homologous chromosomes. Identifiable homologous pairs of chromosomes were taken from respective pictures and displayed according to homeologous pairs between *Quercus* and *Castanea* species and in decreased order of magnitude in *Fagus* **d)**. The Bars= 10 μm

**Figure III.1**

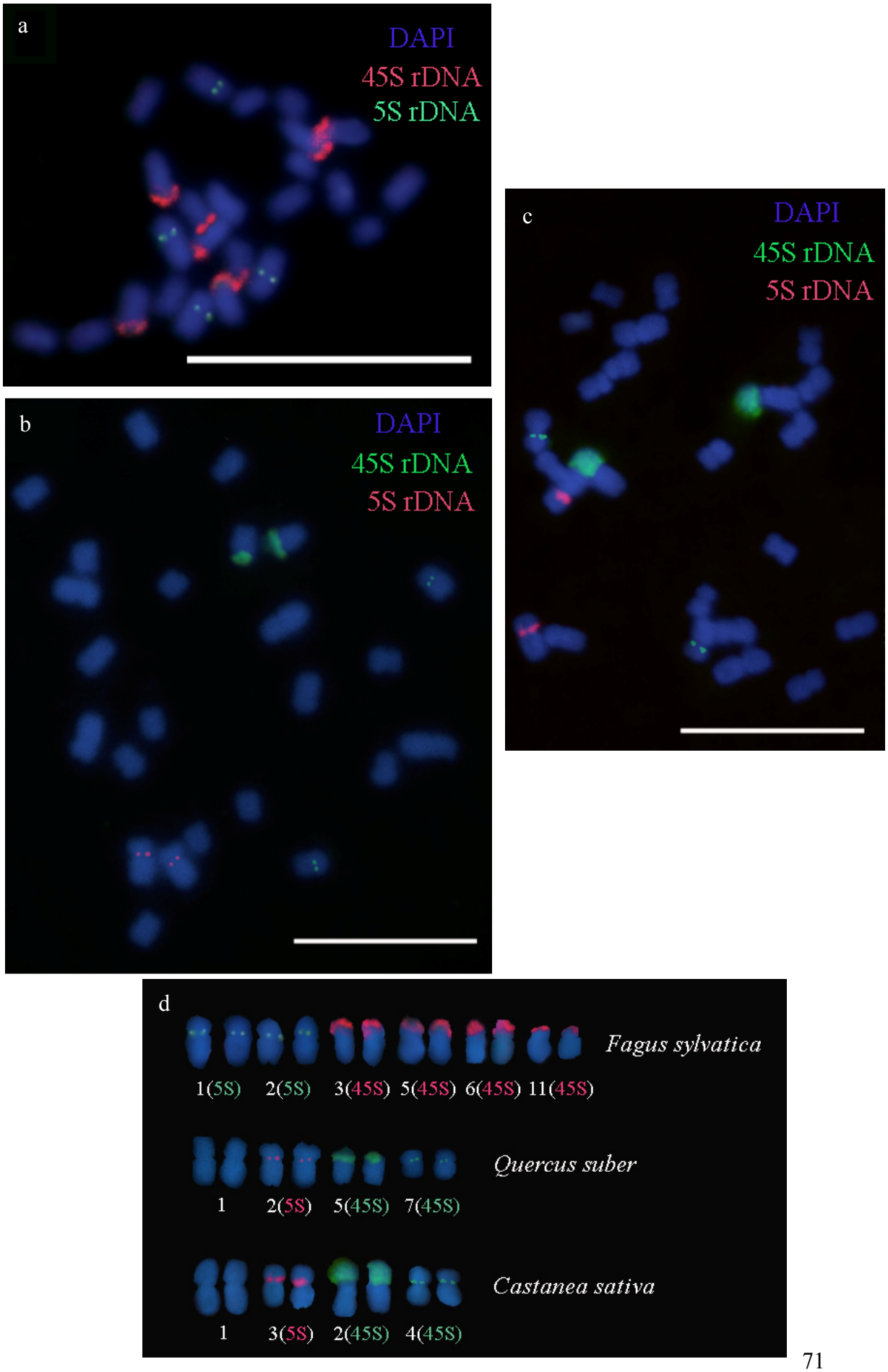

a
DAPI
45S rDNA
5S rDNA
b
DAPI
45S rDNA
5S rDNA
c
DAPI
45S rDNA
5S rDNA
d
Fagus sylvatica
1(5S) 2(5S) 3(45S) 5(45S) 6(45S) 11(45S)
Quercus suber
1 2(5S) 5(45S) 7(45S)
Castanea sativa
1 3(5S) 2(45S) 4(45S)

**Table III.1** Morphometric data from 4 metaphase cells of *Fagus sylvatica*

**a)** Total chromosomes dimension, estimated DNA content per chromosome and ribosomal genes localization.

| Chromosome pair [a] | n (counts) [b] | Total length (μm) Range | Mean±SD | Relative length [c] % | Estimated DNA content [d] (Mbp) | FISH signal 45S | 5S |
|---|---|---|---|---|---|---|---|
| 1 | 7 | 1.81-2.18 | 2.00 ±0.18 | 10.31 | 56.09 | | S |
| 2 | 8 | 1.72-2.04 | 1.85 ±0.22 | 9.54 | 51.90 | | S |
| 3 | 8 | 1.67-1.92 | 1.82 ±0.13 | 9.38 | 51.03 | S | |
| 4 | 7 | 1.65-1.90 | 1.80 ±0.13 | 9.28 | 50.49 | | |
| 5 | 8 | 1.60 -1.88 | 1.75 ±0.13 | 9.02 | 49.07 | S | |
| 6 | 8 | 1.58-1.80 | 1.71 ±0.12 | 8.81 | 47.93 | S | |
| 7 | 7 | 1.48-1.78 | 1.63 ±0.13 | 8.40 | 45.70 | | |
| 8 | 8 | 1.46-1.70 | 1.59 ±0.12 | 8.20 | 44.61 | | |
| 9 | 8 | 1.37-1.64 | 1.49 ±0.11 | 7.68 | 41.78 | | |
| 10 | 8 | 1.13-1.59 | 1.32 ±0.22 | 6.80 | 36.99 | | |
| 11 | 7 | 1.12-1.44 | 1.25 ±0.13 | 6.44 | 35.04 | S | |
| 12 | 8 | 1.07-1.30 | 1.19 ±0.11 | 6.13 | 33.35 | | |

**b)** Lengths of each chromosome arm, centromeric indexes and chromosomes classification

| Chromosome pair [a] | n (counts) | Long arm length (μm) Mean ± SD | Short arm length (μm) Mean ± SD | Centromeric index % [e] Mean ± SD | Chromosome type |
|---|---|---|---|---|---|
| 1 | 7 | 1.12 ± 0.08 | 0.88 ± 0.11 | 44.11 ± 1.61 | m |
| 2 | 8 | 1.05 ± 0.06 | 0.84 ± 0.11 | 45.40 ± 2.51 | m |
| 3 | 8 | 1.13 ± 0.07 | 0.67 ± 0.04 | 36.20 ± 0.38 | sm |
| 4 | 7 | 1.07 ± 0.06 | 0.72 ± 0.07 | 40.23 ± 1.49 | m |
| 5 | 8 | 1.07 ± 0.09 | 0.68 ± 0.04 | 38.95 ± 0.65 | m |
| 6 | 8 | 1.04 ± 0.07 | 0.68 ± 0.04 | 39.60 ± 0.34 | m |
| 7 | 7 | 0.83 ± 0.06 | 0.79 ± 0.08 | 49.01 ± 0.51 | M |
| 8 | 8 | 0.96 ± 0.06 | 0.64 ± 0.06 | 40.14 ± 1.08 | m |
| 9 | 8 | 0.94 ± 0.05 | 0.55 ± 0.06 | 36.91 ± 0.68 | sm |
| 10 | 8 | 0.83 ± 0.13 | 0.49 ± 0.09 | 37.12 ± 0.36 | sm |
| 11 | 7 | 0.94 ± 0.07 | 0.31 ± 0.06 | 24.86 ± 2.31 | st/sm |
| 12 | 8 | 0.66 ± 0.05 | 0.53 ± 0.06 | 44.51 ± 0.99 | m |

SD – Standard deviation
S – Short arm of the chromosome
M – The centromeric position is *sensu stricto* median
m – The centromeric position is in the median region
st – The centromeric position is subterminal

[a] Chromosomes were ordered and numbered according to their rank of the total length at metaphase (full contraction).

[b] Only chromosomes totally visible and not too folded were taken into account.

[c] Relative length = 100(chromosome length/genome length).

[d] Estimated DNA content = relative length x estimated genome size, *i.e.*, 544 Mb (Gallois *et al.*, 1999)

[e] Centromeric index =100 x ( length of the short arm/ total length)

**Table III.2** Morphometric data from 6 metaphase cells of *Quercus suber*

**a)** Total chromosomes dimension, estimated DNA content per chromosome and ribosomal genes localization.

| Chromosome pair [a] | n (counts) [b] | Total length (μm) Range | Mean±SD | Relative length [c] % | Estimated DNA content [d] (Mbp) | FISH signal 45S | 5S |
|---|---|---|---|---|---|---|---|
| 1 | 12 | 2.55-3.32 | 2.96 ±0.23 | 12.12 | 112.84 | | |
| 2 | 12 | 2.35-2.68 | 2.52 ±0.12 | 10.32 | 96.08 | | S |
| 3 | 12 | 2.28-2.60 | 2.44 ±0.13 | 9.99 | 93.01 | | |
| 4 | 12 | 1.88-2.46 | 2.16 ±0.12 | 8.84 | 82.30 | | |
| 5 | 12 | 1.80-2.34 | 2.06 ±0.16 | 8.43 | 78.48 | S | |
| 6 | 12 | 1.72-2.30 | 2.01 ±0.10 | 8.23 | 76.62 | | |
| 7 | 12 | 1.62-2.20 | 1.94 ±0.16 | 7.94 | 73.92 | S | |
| 8 | 12 | 1.62-1.98 | 1.80 ±0.09 | 7.37 | 68.61 | | |
| 9 | 12 | 1.61-1.95 | 1.78 ±0.14 | 7.29 | 67.87 | | |
| 10 | 12 | 1.52-1.88 | 1.70 ±0.11 | 6.96 | 64.80 | | |
| 11 | 12 | 1.46-1.78 | 1.62 ±0.09 | 6.63 | 61.73 | | |
| 12 | 11 | 1.28-1.60 | 1.44 ±0.08 | 5.89 | 54.84 | | |

**b)** Lengths of each chromosome arm, centromeric indexes and chromosomes classification

| Chromosome pair [a] | n (counts) [b] | Long arm length (μm) Mean ± SD | Short arm length (μm) Mean ± SD | Centromeric index % [e] Mean ± SD | Chromosome type |
|---|---|---|---|---|---|
| 1 | 12 | 1.59 ± 0.13 | 1.38 ± 0.12 | 46.02 ± 1.12 | m |
| 2 | 12 | 1.61 ± 0.09 | 0.92 ± 0.05 | 36.51 ± 0.73 | sm |
| 3 | 12 | 1.53 ± 0.10 | 0.90 ± 0.05 | 36.89 ± 0.44 | sm |
| 4 | 12 | 1.44 ± 0.09 | 0.73 ± 0.08 | 33.80 ± 0.32 | sm |
| 5 | 12 | 1.51 ± 0.12 | 0.54 ± 0.09 | 25.01 ± 0.70 | st/sm |
| 6 | 12 | 1.36 ± 0.10 | 0.65 ± 0.09 | 32.34 ± 1.03 | sm |
| 7 | 12 | 1.19 ± 0.08 | 0.75 ± 0.12 | 38.66 ± 1.01 | m |
| 8 | 12 | 1.06 ± 0.08 | 0.70 ± 0.04 | 38.89 ± 0.18 | m |
| 9 | 12 | 1.20 ± 0.10 | 0.59 ± 0.11 | 33.14 ± 0.12 | sm |
| 10 | 12 | 0.86 ± 0.08 | 0.86 ± 0.07 | 50.00 ± 0.08 | M |
| 11 | 12 | 1.03 ± 0.16 | 0.60 ± 0.10 | 37.04 ± 0.18 | sm |
| 12 | 11 | 0.91 ± 0.07 | 0.53 ± 0.02 | 36.81 ± 0.10 | sm |

SD – Standard deviation
S – Short arm of the chromosome
M – The centromeric position is *sensu stricto* median
m – The centromeric position is in the median region
st – The centromeric position is subterminal

[a] Chromosomes were ordered and numbered according to their rank of the total length at metaphase (full contraction).

[b] Only chromosomes totally visible and not too folded were taken into account.

[c] Relative length = 100 x (chromosome length/genome length).

[d] Estimated DNA content = relative length x estimated genome size, *i.e.*, 931 Mb (Zoldoš *et al*., 1998)

[e] Centromeric index =100 x ( length of the short arm/ total length)

**Table III.3** Morphometric data from 8 metaphase cells of *Castanea sativa*

**a)** Total chromosomes dimension, estimated DNA content per chromosome and ribosomal genes localization.

| Chromosome pair [a] | n (counts) [b] | Total length (μm) Range | Total length (μm) Mean±SD | Relative length [c] % | Estimated DNA content [d] (Mbp) | FISH signal 45S | FISH signal 5S |
|---|---|---|---|---|---|---|---|
| 1 | 14 | 1.84-3.73 | 3.05 ±0.60 | 11.49 | 109.84 | | |
| 2 | 13 | 1.58-3.68 | 2.94 ±0.64 | 11.08 | 105.92 | S | |
| 3 | 11 | 1.55-3.34 | 2.69 ±0.55 | 10.14 | 96.93 | | S |
| 4 | 13 | 1.47-3.07 | 2.39 ±0.48 | 9.01 | 86.14 | S | |
| 5 | 10 | 1.35-2.83 | 2.19 ±0.50 | 8.25 | 78.87 | | |
| 6 | 10 | 1.30-2.61 | 2.06 ±0.48 | 7.76 | 74.19 | | |
| 7 | 9 | 1.28-2.33 | 2.00 ±0.43 | 7.54 | 72.08 | | |
| 8 | 10 | 1.18-2.38 | 1.98 ±0.41 | 7.46 | 71.32 | | |
| 9 | 7 | 1.13-2.35 | 1.89 ±0.46 | 7.12 | 68.07 | | |
| 10 | 16 | 1.11-2.36 | 1.84 ±0.42 | 6.94 | 66.35 | | |
| 11 | 14 | 1.12-2.26 | 1.81 ±0.36 | 6.82 | 65.20 | | |
| 12 | 12 | 1.07-2.13 | 1.69 ±0.36 | 6.37 | 60.90 | | |

**b)** Lengths of each chromosome arm, centromeric indexes and chromosomes classification

| Chromosome pair [a] | n (counts) | Long arm length (μm) Mean ± SD | Short arm length (μm) Mean ± SD | Centromeric index % [e] Mean ± SD | Chromosome type |
|---|---|---|---|---|---|
| 1 | 14 | 1.59 ± 0.28 | 1.47 ± 0.32 | 47.85 ± 1.57 | m |
| 2 | 13 | 1.57 ± 0.35 | 1.33 ± 0.34 | 45.61 ± 2.83 | m |
| 3 | 11 | 1.51 ± 0.31 | 1.12 ± 0.26 | 42.32 ± 2.31 | m |
| 4 | 13 | 1.34 ± 0.29 | 1.07 ± 0.24 | 44.26 ± 2.12 | m |
| 5 | 10 | 1.32 ± 0.31 | 0.87 ± 0.20 | 39.69 ± 1.16 | m |
| 6 | 10 | 1.16 ± 0.28 | 0.90 ± 0.21 | 43.73 ± 1.04 | m |
| 7 | 9 | 1.01 ± 0.22 | 0.99 ± 0.21 | 49.50 ± 0.71 | M |
| 8 | 10 | 1.01 ± 0.21 | 0.98 ± 0.20 | 49.16 ± 0.82 | m/M |
| 9 | 7 | 1.14 ± 0.25 | 0.74 ± 0.22 | 40.09 ± 2.14 | m |
| 10 | 16 | 0.99 ± 0.22 | 0.85 ± 0.18 | 46.44 ± 1.37 | m |
| 11 | 14 | 1.12 ± 0.24 | 0.67 ± 0.14 | 36.41 ± 3.55 | sm/m |
| 12 | 12 | 0.91 ± 0.22 | 0.78 ± 0.15 | 46.46 ± 2.36 | m |

SD – Standard deviation
S – Short arm of the chromosome
M – The centromeric position is *sensu stricto* median
m – The centromeric position is in the median region

[a] Chromosomes were ordered and numbered according to their rank of the total length at metaphase (full contraction).

[b] Only chromosomes totally visible and not too folded were taken into account.

[c] Relative length = 100(chromosome length/genome length).

[d] Estimated DNA content = relative length x estimated genome size, *i.e.*, 956 Mb (Barow and Meister, 2002)

[e] Centromeric index =100( length of the short arm/ total length)

**Table III.4**- Chi Square values, degree of freedom (df) and p value for the karyotype distribution of chromosomes through chromosome morphometric classes of *Fagus sylvativa*, *Quercus suber* and *Castanea sativa*

| Species analysed | Chi-Square | df | p |
|---|---|---|---|
| *F. sylvatica, Q. suber and C. sativa* | 9.54 | 6 | 0.1454 |

**Table III.5** – Karyotype asymmetry index (AI), calculation using the coefficient of variation of chromosome length ($CV_{CL}$) and the coefficient of variation of centromeric index ($CV_{CI}$). *S* stands for the standard deviation and *X* for the mean

| Species | $CV_{CL}$ | $CV_{CI}$ | AI |
|---|---|---|---|
| *Fagus sylvatica* | 16.05 | 13.36 | 2.14 |
| *Quercus suber* | 21.08 | 17.08 | 3.60 |
| *Castanea sativa* | 20.64 | 9.12 | 1.88 |

$$CV_{CL} = \frac{S_{CL}}{X_{CL}} * 100 \qquad CV_{CI} = \frac{S_{CI}}{X_{CI}} * 100 \qquad AI = \frac{CV_{CL} * CV_{CI}}{100}$$

**Figure III.2** – Fluorescent in situ hybridization in c-metaphase of *Castanea sativa* Mill. **a)**, prometaphase of *Quercus suber* L. **b)** and c-metaphase of *Fagus sylvatica* L. **c)** using the Arabidopsis-like telomere sequence as FISH probe (green signals). DNA counterstained with DAPI. Bars= 10 μm.

**Figure III.3** – Ideograms of Fagaceae species based on values of tables 1,2,3. The chromosomes are arranged according to decreasing length. The 5S rDNA *loci* are represented in green while the 45S rDNA *loci* are represented in red in all species.

**Figure III.2**

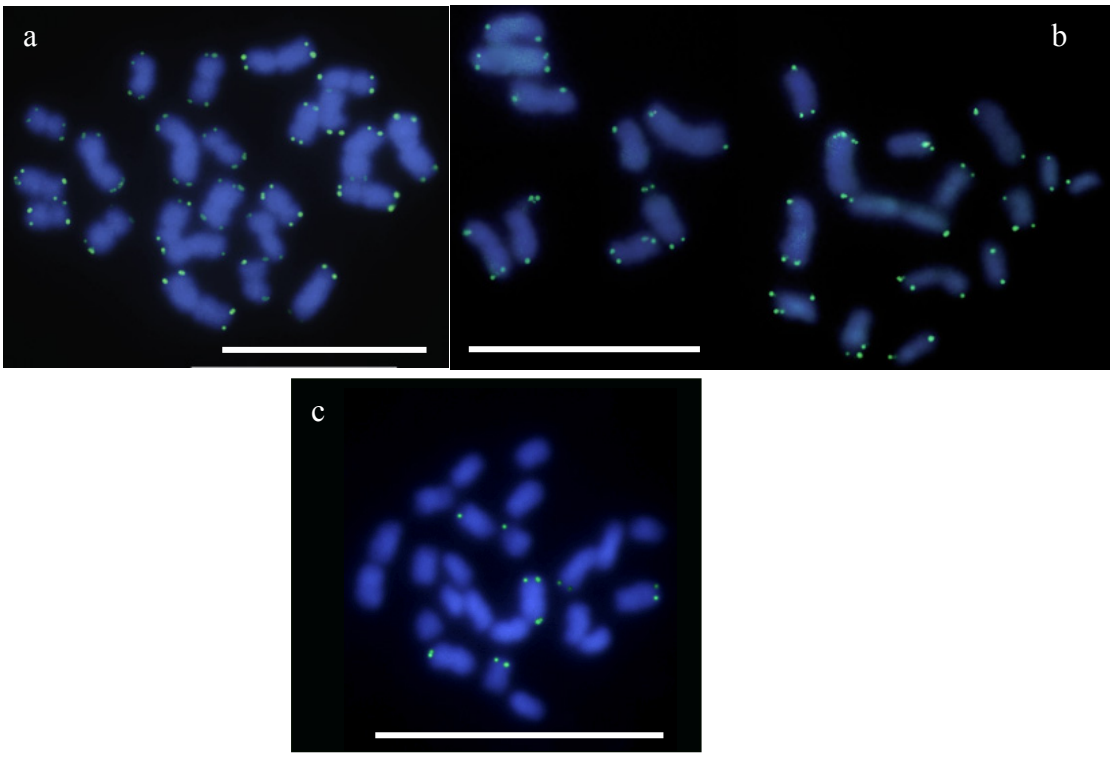
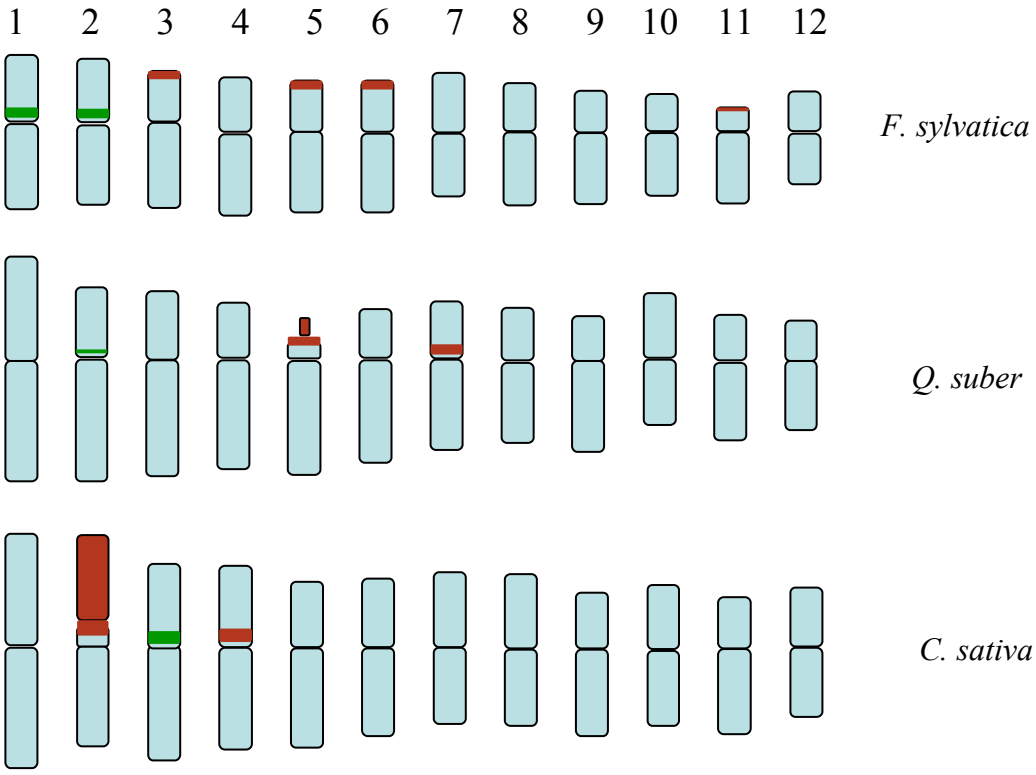



**Figure III.3**



### III.3.2. Intraspecific genome variability within Quercus suber

During this work it was detected one individual of *Quercus suber* presenting 36 chromosomes (Fig. III.4 a).

Morphological analysis performed with DAPI staining revealed that each chromosome have three representatives (Fig. III.4 a arrows, and b). FISH with rDNA markers on these oak chromosomes revealed three copies of each *locus*: three chromosomes bearing major NORs, three chromosomes with minor NORs, and three with 5S *loci*. This data support the triploid nature of that plant, which seems to have an extra complete haploid complement set of chromosomes, 2n=3x=36. All the rDNA signals are on the expected positions indicating that no major rearrangements occurred in this plant (Fig. III.4 a).

#### III.3.2.9. 5S rDNA organization in pollen nuclei

*Q. suber* mature pollen grains collected from trees grown in Sintra, Portugal, analysed under the bright field microscope without cover slip revealed similar dimensions (Fig. III.5 a). Cork oak mature pollen nuclei were isolated from pollen grains and analysed with DAPI staining. Two morphologically different populations were observed: generative nuclei with a condensed chromatin organization and vegetative nuclei with more decondensed chromatin (Fig. VI.1). In order to find unreduced pollen grains, FISH with 5S rDNA Qsu5S probe was performed in these two populations and the number and organization of these *loci* was evaluated in generative nuclei due to its more condensed nature. Within a population of 250 generative nuclei 67.2 % showed one FISH signal (Fig. III.5 b), 32.8 % presented one split FISH signal (Fig. III.5 c) or two closely adjacent signals (Fig. III.5 d). The single signal (Fig. III.5 b) is bigger than individual signals present in two signals cells (Fig. III.5 c,d) probably representing juxtaposition of sister chromatids.

### III.3.3. Cytogenetic analysis of European and Asian species of genera Quercus and Castanea

Meristematic root tip cells from the European *Quercus ilex* subsp. *rotundifolia* (Fig. III.6 a), the Asian *Quercus acutissima* (Fig. III.6 b) and *Quercus serrata* (Fig. III.6 c), were analysed and revealed 24 chromosomes (2n=24) (Fig. III.6 a,b,c) as the other European and American *Quercus* species described in Zoldoš *et al*. (1999).

While the other species lack evident heterochromatic blocks, the analysis of prometaphase chromosomes showed DAPI positive spots at the centromere region of *Q. serrata* (Fig. III.6 c inset). rDNA mapping of 5S and 45S sequences on these *Quercus* species showed the typical pattern of *Quercus* species (Zoldoš *et al*., 1999): one 5S rDNA pericentromeric *locus*, one terminal and bigger 45S rDNA *locus* (major NOR) and a pericentromeric and smaller 45S rDNA *locus* (minor NOR). Considering the intensity of FISH label all *Quercus* species analysed revealed more 5S and 45S rDNA copies than *Quercus suber* (Fig. III.1 b). The remaining karyotype is very similar to the *Quercus suber* karyotype previously described; being the most distinctive chromosome pair the biggest metacentric one.

In the genus *Castanea*: chromosomes from the Asian species, *C. crenata* (Japanese chestnut) (Fig. III.6 d) and *C. mollissima* (Chinese chestnut) (Fig. III.6 e) were analysed and revealed that these species also have 24 chromosomes (2n=24) as described in Jaynes (1962).

FISH with rDNA probes revealed that *C. crenata* (Fig. III.6 d) has the same *loci* number and physical location as *C. sativa* (Fig. III.1 c), while *C. molissima* shows differences in relation to the location of the major NOR, and on the number of 5S *loci*. The Chinese chestnut shows two 45S rDNA *loci* and one 5S *loci* in a pericentromeric position. One chromosome pair comprises the major 45S NOR and a 5S rDNA *locus* located towards the telomere. (Fig. III.6e and 6f). The remaining chromosomes compose similar karyotypes where the notorious chromosome is the biggest metacentric one. (Fig. III.6 f).

**Figure III.4** – Fluorescent in situ hybridization in c-metaphase of a triploid individual of *Quercus suber* L. 2n=3x=36, using the ribosomal DNA probes (pTa71 and 5SQsu) as FISH markers in helping detecting homologous chromosomes. The four most easily detectable triplets of homologous chromosomes are evidenced by arrows with the same colour in **a)**. Bar= 10 μm. Karyotype of the same c- metaphase has triplets of homologs **b)** arranged according the *Quercus suber* ideogram in Fig. III.3.

**Figure III.5** – Analysis of pollen grains from *Quercus suber* trees. **a)** Screening in bright field microscope of mature pollen grain **b,c,d)** FISH analysis on the generative nucleus using the 5S rDNA probe specific to *Q. suber* (green), showing one signal **(b)** or two partial superimposed signals **(c)** and two adjacent signals **(d)**, corresponding to the two chromatids. Bar=50 μm in **a)**. Bars=10 μm in **b,c,d)**.

**Figure III.4**

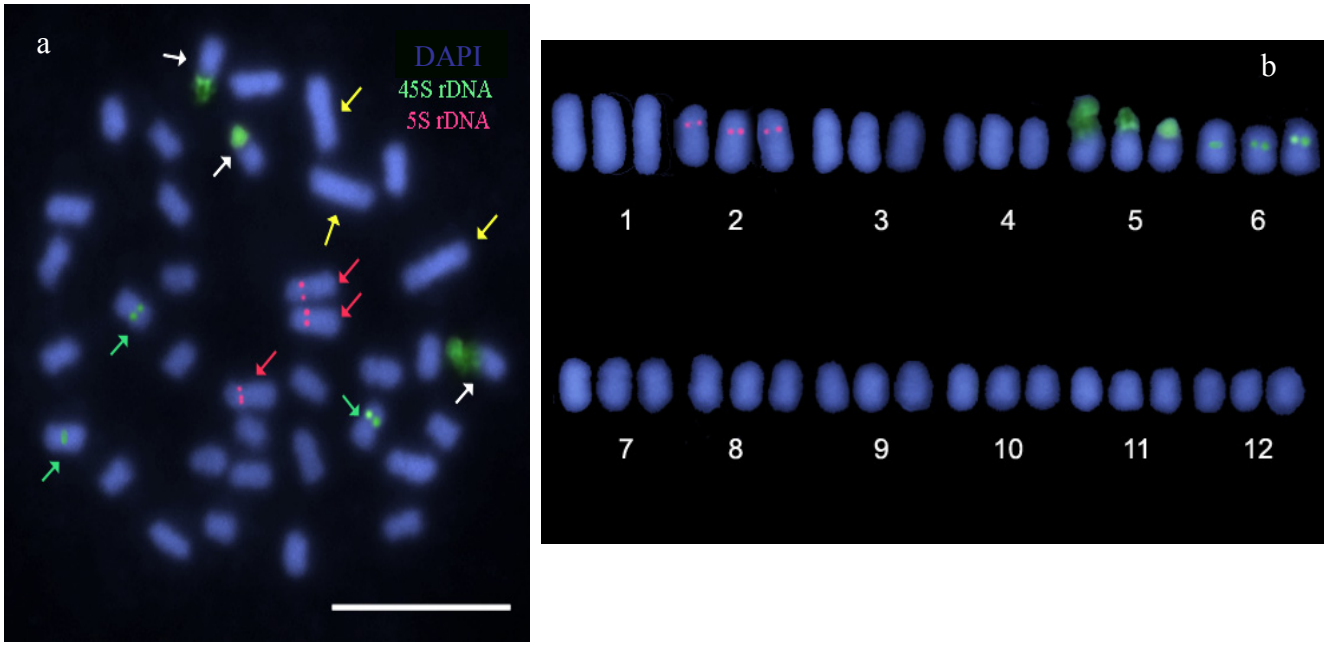



**Figure III.5**

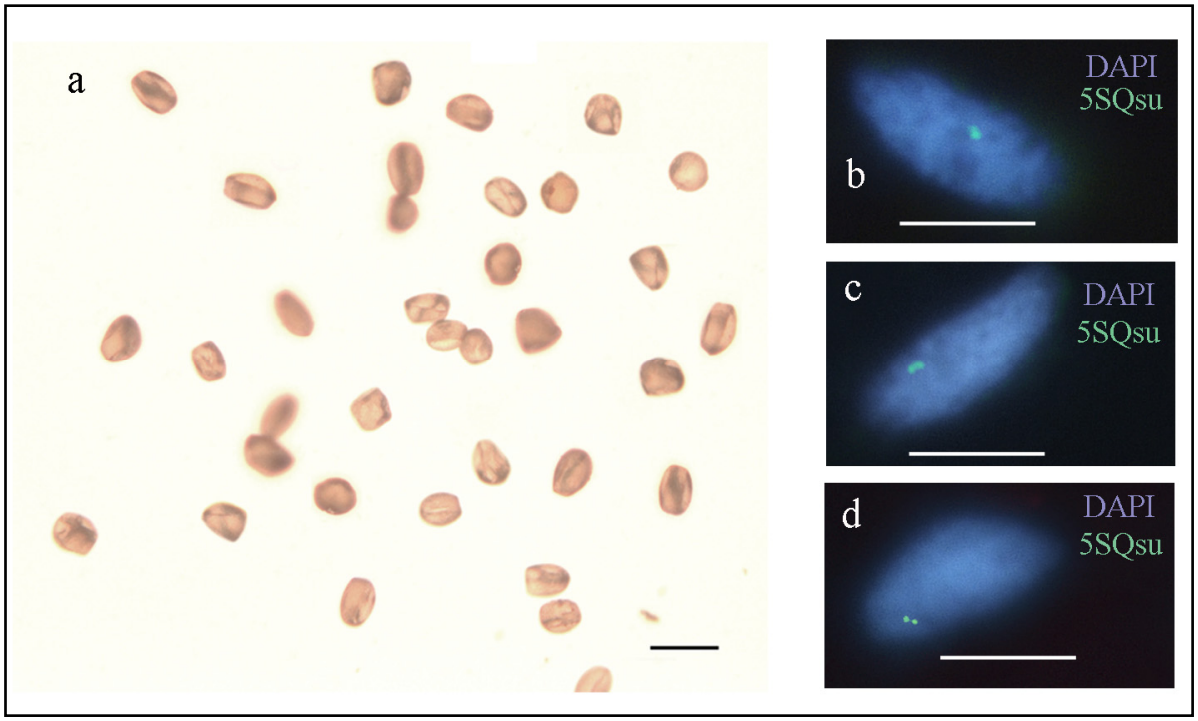
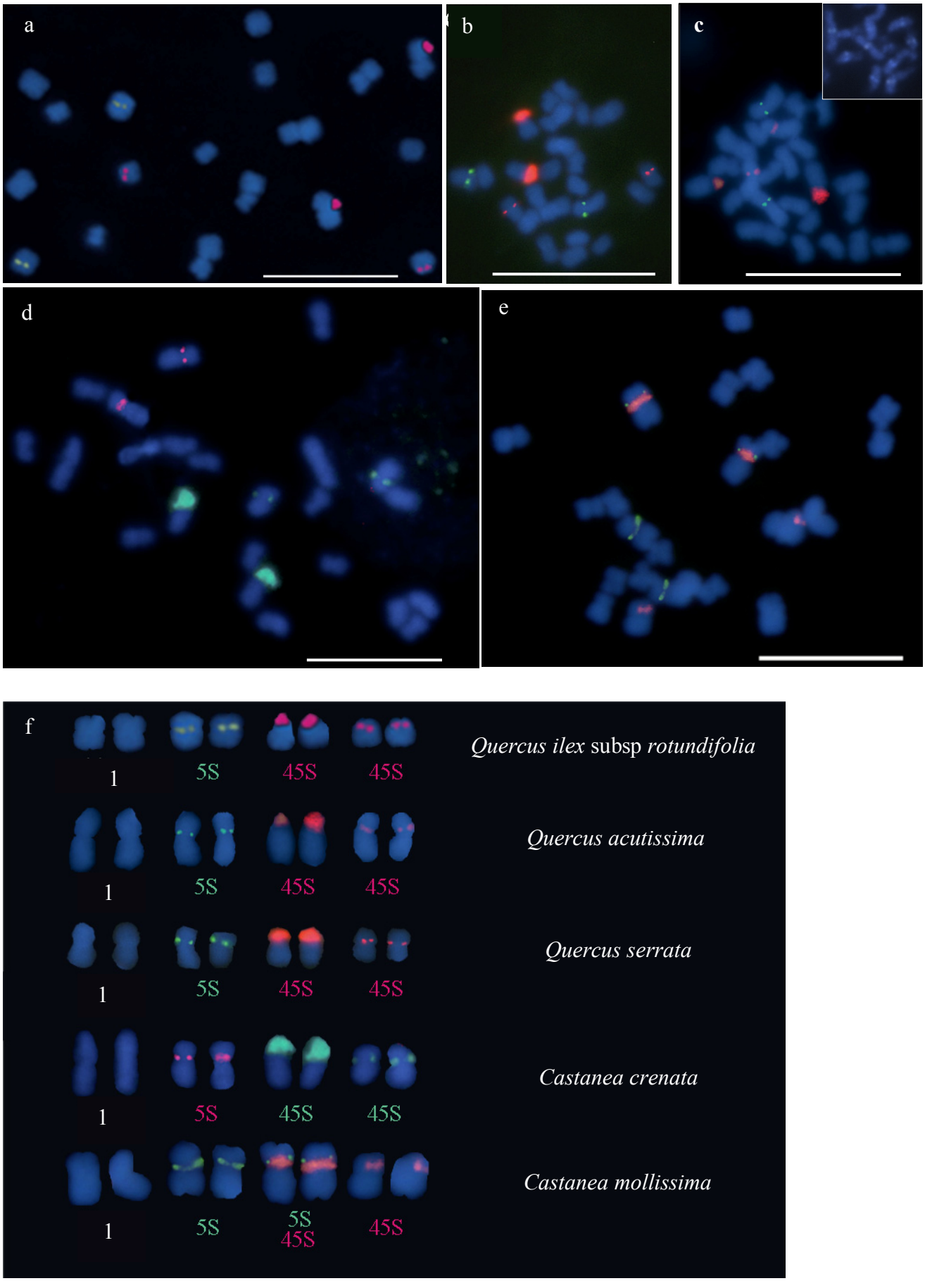


**Figure III.6** – Fluorescent in situ hybridization on c-metaphases of *Quercus ilex* L. subsp *rotundifolia* (Lam) O. Schwarz **a)**, *Quercus acutissima* Carruth. **b)**, *Quercus serrata* Murray **c)**, *Castanea crenata* Sieb. & Zucc **d)**, *Castanea mollissima* Bl. **e)**, with the 45S and 5S ribosomal DNA probes. Prometaphase chromosomes of *Q. serrata* showing DAPI positive spots on the centromere region **c)** inset. Identifiable homologous pairs of chromosomes were taken from respective pictures and organized between species accordingly to homeologous pairs **f)**. Bars= 10 μm.

**Figure III.6**


a
b
c
d
e
f
1
5S
45S
45S
Quercus ilex subsp rotundifolia
Quercus acutissima
Quercus serrata
Castanea crenata
Castanea mollissima

## III.4. Discussion

The karyotypes obtained for the three Fagaceae species, according to the morphometric classification of Levan *et al.* (1964) are: 1M, 7m, 3sm, 1st/sm for *Fagus sylvatica*; 1M, 3m, 7sm, 1 st/sm for *Quercus suber*, and 1M, 9m, 1m/M, 1sm/m for *Castanea sativa*. The variability present in the measurements turned impossible to clearly assign some of the chromosomes to specific categories. Using the same chromosome classification *Fagus sylvatica* karyotype has been already characterized as 3M, 3m, 3sm and 3st **(**Ohri and Ahuja, 1991), and *Quercus suber* karyotype, was characterized as 7m and 5sm (D'Emerico *et al*., 2000), or with one st pair (Zoldoš *et al*., 1999). The karyotype of *Quercus suber* obtained in this work, although closer to the karyotype described by Zoldoš *et al*. (1999) which was also based on rDNA FISH markers, describes a Metacentric *sensu stricto* (M) pair that was not described in that previous study. The differences between our *F. sylvatica* and *Q. suber* karyotypes and those described by other authors can be related with measurements performed on different types of chromosomes. Differential condensation of chromatin fibres at prometaphase affects chromosomes pair's measurements as has been described in *Brassica* (Nakayama and Fukui, 1997) and *Quercus* (Ohri and Ahuja, 1990; Zoldoš *et al*., 1999). Also for molecular size estimation, it is important to target metaphase chromosomes where is more uniform along the chromosome long axis and relative lengths most accurately reflecting relative molecular size (Kim *et al*., 2005). Although the contribution of each chromosome is similar in the three karyotypes, varying in average from six to twelve percent, the amount of DNA of each chromosome is very different: for instance the biggest chromosome of *Fagus sylvatica* has less DNA than the smallest of *Castanea sativa* (Table III.1 a and Table III.3 a) and little more than the smallest of *Quercus suber* (Table III.2 a). The evaluation of karyotype asymmetry revealed by the asymmetric index show that *Quercus suber* karyotype is the most asymmetric (AI= 3.60), followed by *Fagus sylvatica* with AI = 2.14 and finally the most symmetric is *Castanea sativa* (AI=1,88). The level of karyotype asymmetry is usually used for the elucidation of phylogenetic relationships (reviewed in Paszko, 2006.). It is also accepted that karyotype evolution generally tends towards an increasing number of acrocentric chromosomes, thereby minimising the risk of deleterious rearrangements. And the opposite tendency, the formation of metacentric chromosomes, is considered to

be the result of 'rare back-eddies' that are generated at random and tolerated or even favoured when they provide short-term advantages (reviewed in Shubert, 2007). From the three genera studied, *Castanea* is considered more recent than the others (Manos *et al*., 2001; Li *et al*., 2004), which according to this definition should be regarded as a result of 'rare back-eddies' in the evolution of Fagaceae.

Although the *Quercus* and *Castanea* genera diverged ~60 million years ago (Manos and Steele, 1997), a very high level of genome conservation has been detected between *Quercus robur* and *Castanea sativa*, using 55 orthologous molecular markers, including expressed sequence tag (EST)-derived markers, SSR markers and 5S rDNA (Casasoli *et al*., 2006). This study also suggests that the 12 linkage groups are homeologous, showing conserved macrosynteny and macrocolinearity, and that no major rearrangements were observed. Our results can also assign with certainty four homeologous chromosomes between *Quercus suber* and *Castanea sativa*: the biggest metacentric and the three chromosomes bearing rDNA landmarks (Fig. III.1 d) although occupying distinct karyotypes positions. Amplification of cluster satellite DNA may differentially alter chromosome morphology without changing the order of other sequences (Navratilova *et al*., 2003), and may explain the differences in the karyomorphology. Our work undoubted recognise *Q. suber* chromosome 2 similar to *C. sativa* chromosome 3 corresponding to the homeologous chromosomes Q5 and C4 in the comparative linkage maps presented in Casasoli *et al*. (2006); and chromosomes 5 and 7 of *Q. suber* similar to 2 and 4 of *C. sativa*, respectively. Chi square analysis of the karyotypes distributions revealed that there are no significant differences between the three Fagaceae karyotypes. Ribosomal *loci* in the karyotypes of these three genera can be correlated with their phylogenetic relationships, (i.e) the early divergence of *Fagus* in the evolution of Fagaceae and the closeness of *Castanea* and *Quercus* genera, proposed by Manos *et al*. (2001) and Li *et al*. (2004), based on DNA sequences, including the 5.8S rRNA gene and two flanking internal transcribed spacers (ITS).

All the Fagaceae species analysed have the same chromosome number, 2n=24, and the differences in the karyotypes seem improbable to be due to major chromosomes rearrangements like centric fusions. However, inversions may have occurred and could be responsible for instance for the presence of pericentromeric 45S rDNA in *Quercus* and *Castanea*, while in *F. sylvatica* all the 45SrDNA *loci* are terminal. FISH with the telomeric sequence failed to reveal macrorearrangements as was observed in *Arabidopsis thaliana* (Uchida *et al*., 2002) since only terminal signals were detected.

The amount of telomeric sequences revealed by the intensity and number of signals detected shows great variability in the species analysed: in *Castanea sativa* all chromosomes ends showed hybridization, while in *Q. suber* and *F. sylvatica* the number and intensity decreased. A direct correlation between the amount of telomeric sequences and genome size can be detected since *C. sativa* has the biggest genome (1C= 956 Mb) and shows a high number and strong signals, while *F. sylvatica* with the smallest genome (1C= 544 Mb), only presents some telomeric signals. The contribution of telomeric DNA to increase genome size is emphasized by the correlation detected in several plant species: in the small genome species Arabidopsis telomeres correspond to 2 to 5 kb, and in large genome species they can contribute up to 12 to 15 kb in cereals and to 60 to 160 kb in tobacco (reviewed in Heslop-Harrison, 2000).

Besides the high degree of conservation of the *Quercus* karyotype, other kind of diversity can be detected within *Quercus* species. Spontaneous intraspecific triploid individuals have been reported in few *Quercus* species (Burda and Shchepotiev, 1973; Butorina, 1993; Dzialuk *et al*., 2007), and here we also report intraspecific triploidy in *Quercus suber*. At least, three mechanisms are described that can lead to intraspecific triploidy: chromosomal non reduction during microsporogenesis or megasporogenesis, formation of 2n pollen grains trough parallel spindles instead of perpendicular ones in the second meiotic division (Butorina, 1993; Zhang *et al*., 2007), and absence of the second meiotic division (Mok and Peloquin, 1973). Screening of *Quercus suber* mature pollen failed to reveal differences in pollen sizes as already detected in *Populus tomentosa* due to non reduced pollen grains (Zhang *et al*., 2007). Evaluation of pollen nuclei ploidy was performed with FISH in generative nuclei due to their highly condensed chromatin using the 5S rDNA probe due to the interphase condensed behaviour of these DNA sequences (reviewed in Ribeiro *et al*., 2008). This analysis shows one bigger or two small and side by side signals corresponding to two chromatids what confirms the G2 nature of these nuclei (Bino *et al*., 1990) but failed to reveal diploid nuclei. As the production of 2n pollen can be dependent on weather conditions as was observed for Chinese white poplar in different years (Kang, 2002 (reviewed in Zhang *et al*., 2007) the pollen of *Q. suber* analysed could not have been affected by that disturbing influence. During pollen germination spontaneous sperm nuclei fusion is normally prevented but special *in vitro* conditions make tobacco sperm cells strongly fusigenic (Tian and Russell, 1998). *Q. suber* as all Mediterranean species suffer high temperatures together with drought during late spring. We can speculate that during the

development and delivery of the sperm cells to the female gametophyte, which occurs during late spring time, some abnormal conditions such as high temperatures and dehydration can induce chemical alterations that may cause unusual sperm fusion (2n) leading to triploid embryos. Finally, polyembryony is a common feature of oak seeds and has already been proposed as a mechanism for *Quercus robur* triploids formation (Johnsson, 1946; Burda and Schepotiev, 1973). In this circumstances the fusion of female or male gametes, will produce, during fertilization, a triploid embryo.

### III.4.1. Proposed models for karyotypes evolution in Fagaceae

In order to understand karyotype evolution in the Fagaceae family, several species belonging to the two genera *Quercus* and *Castanea* with very distinct geographical distributions were studied: the Mediterranean *Quercus ilex* subsp *rotundifolia*, and the Asian species, *Quercus acutissima* and *Quercus serrata* all belonging to the subgenus *Quercus*, as well as two Asian chestnut species, the Chinese chestnut, *Castanea mollissima*, and Japanese chestnut, *Castanea crenata*, belonging to the section *Eucastanon* as well as *Castanea sativa.*

Analysis of chromosome and rDNA *loci* numbers and location agree with the highly homogeneous nature of *Quercus* genus. The Asian species were the missing link that allows us to say with certainty that the subgenus *Quercus* have a highly conserved karyotype, independent of their distinct geographical distributions. Despite the high level of karyotype conservation, *Quercus serrata* was the only species in this study presenting distinct centromeric DAPI positive bands as already described for some other *Quercus* spp. in Zoldoš *et al*. (1999).

Contrasting with *Quercus*, the genus *Castanea* shows some divergence since the two Asian chestnuts display different rDNA *loci* number and location: *C. mollissima*, the Chinese chestnut, present the major NOR located interstitially, and shows an extra 5S rDNA *locus* adjacent to the major NOR and located distally to the centromere while the other Asian species, *C. crenata* (Japanese chestnut), share the same number and location of rDNA *loci* with the European chestnut.

These data suggest that *Quercus* karyotype seem to have an evolutionary conservation while *Castanea* karyotypes have different adaptations regarding its

geographical distribution. Differential amplification of repetitive sequences either coding or non-coding can account for karyotypes differences. The 45S rDNA and to a less extent 5S rDNA positions are highly polymorphic and well known for their potential intragenomic mobility being unequal recombination, transposition and chromosomal rearrangements such as inversions, insertions/deletions and translocations pointed as responsible for this mobility (reviewed in Schubert, 2007).

It is remarkable that *Fagus sylvatica* which has the smallest genome, and which genus is pointed to be basal in the phylogeny of Fagaceae has the highest number of rDNA *loci*. Analysis of FISH intensity signals revealed that *F. sylvatica* has more rDNA copies than *Quercus suber*, although contrasting with the positive correlation between rDNA copy number and genome size that have been proposed by Prokopowich *et al*. (2003).

During Fagaceae evolution several chromosomal rearrangements not affecting chromosome number may have been involved in the reduction of rDNA *loci*. A model to explain the evolution from an ancestral karyotype of the genus *Fagus* to the genus *Quercus* is proposed in Fig III.7. The reduction of 5S *loci* (from two in *Fagus* to one in *Quercus*) can be explained by an interstitial chromosome deletion involving the pericentromeric 5S rDNA *locus* **(a)**. Indeed, in animals and plants, the pericentromeric domains have long been recognized as highly dynamic chromosomes regions with extensive level of rearrangements including deletions and inversions (Eichler and Sankoff, 2003).

From the four terminal 45S rDNA *loci* of *Fagus* only one is detected in *Quercus*, therefore a relocation of at least one terminal 45S rDNA *locus* of *Fagus* had to occur during *Quercus* evolution. A paracentric inversion involving one whole arm would transform the terminal NOR into a pericentric one **(b)** while several mechanisms such as amplification/deletion of the remaining three terminal 45S rDNA could be pointed for this reduction. Indeed, meiotic misalignment and interlocus unequal crossing over involving repetitive sequences of non-homologous chromosomes have been pointed to produce duplications and deletions of chromosome segments, particularly for terminal 45S rDNA loci **(c)** (Arnheim *et al*., 1980; Wendel *et al*., 1995; Hanson *et al*., 1996; Pedrosa-Harand *et al*., 2006). Non-reciprocal translocation (**d**) is another mechanism suggested to achieve sequence amplification as already proposed (King, 2002). Global DNA amplification affecting equally **(e)** or differentially **(f)** both chromosomes arms is the final mechanism proposed to increase genome size. Much of the rearrangements

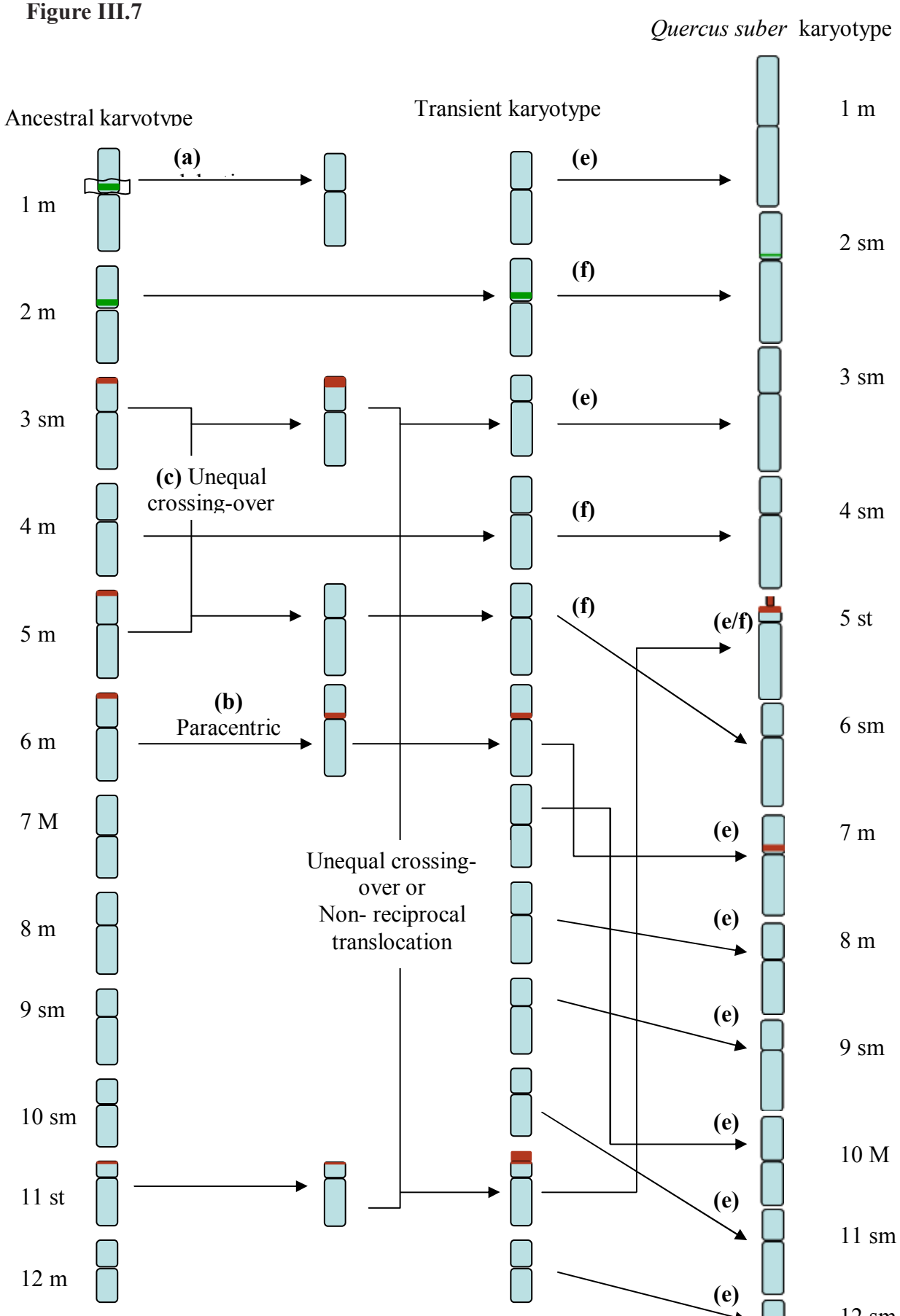
proposed can be related with transposition events which due to chromosome-breaking, aborted transposition or ectopic recombination between homologous transposable elements at different chromosomal locations can stimulate some types of genome rearrangements like inversion, duplication or deletion of adjacent DNA (reviewed at Bennetzen, 2005).

DNA sequence data from different regions of the chloroplast genome (Lang *et al*., 2006) point to the genus *Castanea* as a monophyletic group with *Castanea crenata* in a basal position followed by a diversification within Asia and then a dispersion and divergence between the Chinese (*Castanea mollissima*) and the European (*Castanea sativa*)/North American species. However *C. crenata* and *C. sativa* share the number and mapping of rDNA *loci* that differs from *C. mollissima*. In this species the 45S rDNA has an intercalary location, and an extra 5S rDNA *locus* is found at the short arm. Whether the fixation of these chromosomal rearrangements occurred during the initial divergence of *C. mollissima* or after speciation is not known, but it seems more plausible that it took place long after speciation, at least after divergence with *C. sativa*. A model to explain the potential mechanisms responsible for the rearrangements that could have happened during the evolution of a *Castanea crenata* karyotype-like to the present karyotype of *Castanea mollissima* is proposed in Fig III.8. A paracentric inversion **(a)** could have occurred to generate the contemporary pattern of *Castanea mollissima.* The transposition **(b)** of 5S rDNA small numbers of copies may have happened by the activity of class II mobile elements. In fact the *Enhancer/Supressor-mutator* (*En/Spm)* transposons family often contains sequences similar to the 5S rDNA genes (Wicker *et al*., 2003), and clusters of the *En/Spm* transposon were found with additional 5S rDNA clusters in *Aegilops*, suggesting that the transfer of 5S rDNA was mediated by these transposons (Raskina *et al*., 2004 a,b). It has also been suggested that part of variability and rapid rearrangement of 5S rDNA *loci* could be attributed to the transposition via *Cassandra* retrotransposons. *Cassandra* elements belong to class I of very special mobile elements, the terminal repeat retrotransposons in miniature (TRIMs) (Kalendar *et al*., 2008). The TRIM elements are composed of 100- to 250-bp LTRs, priming sites for reverse transcriptase, a small intervening segment and completely lack reading frames for retrotransposon proteins. Thus TRIM elements are not capable of autonomous transposition and probably require the help of mobility-related proteins encoded by other retrotransposons (Havecker *et al*., 2004). The Cassandra TRIMs in all species investigated carry conserved 5S RNA sequences and associated RNA

polymerase (pol) III promoters and terminators in their long terminal repeats (LTRs) suggesting that the maintenance of the 5S RNA domain may aid *Cassandra* replication (Kalendar *et al*., 2008). Finally transposons or other repetitive elements may be involved in the amplification of the chromosome satellite **(c)**.

**Figure III.7** – Model for the possible evolution events explaining the transition from an ancestral karyotype of genus *Fagus* to the genus *Quercus*

**Figure III.7**

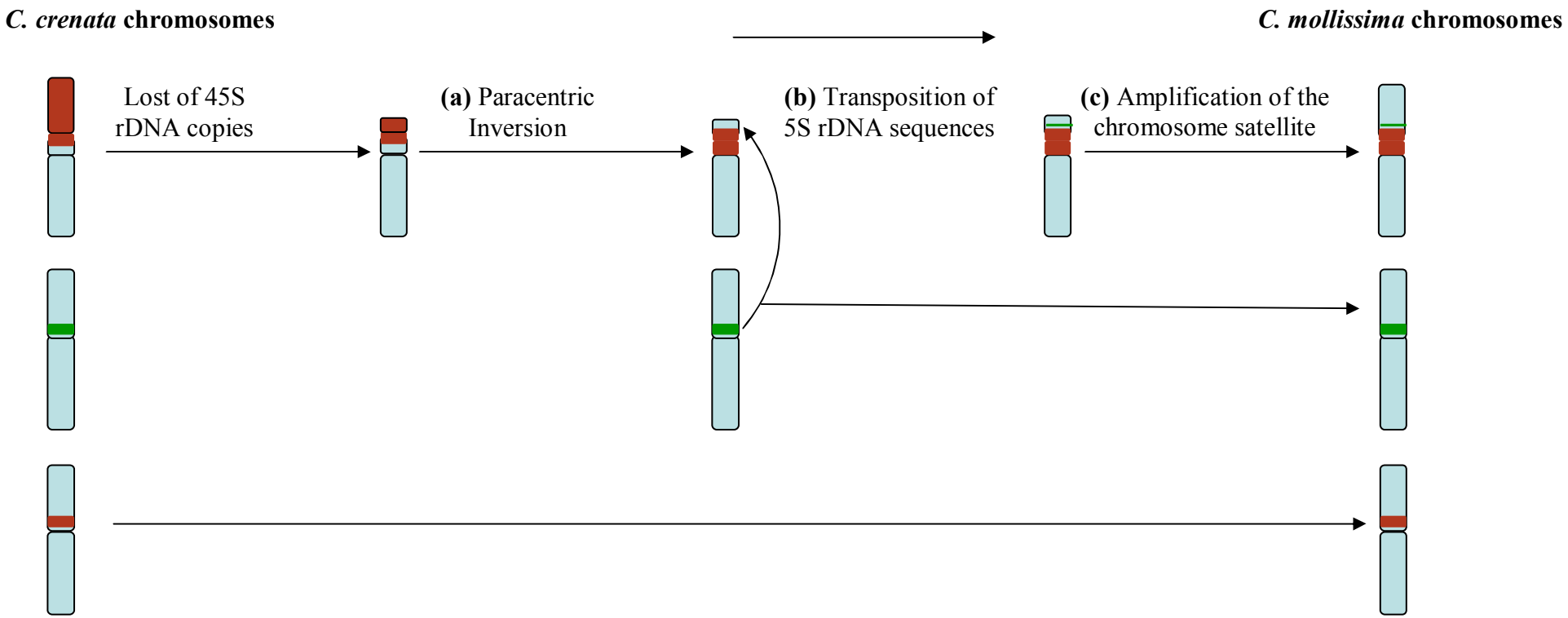



(e) – Equal DNA amplification

(f) –Diferential DNA amplification

**Figure III.8** – A model for possible rearrangements involving two chromosomes with rDNA sequences during the evolution of *C. crenata* to *C. mollissima* karyotype

**Figure III.8**


*C. crenata* chromosomes
*C. mollissima* chromosomes
Lost of 45S rDNA copies
**(a)** Paracentric Inversion
**(b)** Transposition of 5S rDNA sequences
**(c)** Amplification of the chromosome satellite

# *IV. Comparative analysis of the genomic functional domain – NOR*

## IV.5. Introduction

In plants nucleolar organizer regions (NORs) are composed of thousands of 45S rRNA genes and are characterized by different structural and functional domains. These domains can present a high degree of condensation or can be decondensed in interphase nuclei and in the metaphase secondary constrictions where they are approximately tenfold less condensed than the adjacent chromatin (Heliot *et al*., 1997). This high level of decondensation results in reduced DNA dye binding, giving rise to the apparent gap in mitotic nucleolar chromosomes. Scanning electron microscopic investigations of plant NORs have shown that these secondary constrictions are indeed structurally different from surrounding chromatin being characterized by parallel fibrils with sparse DNA distribution (Wanner and Formanek, 1995). In addition, secondary constrictions are docking sites for nucleolar proteins involved in nucleolus formation and rDNA transcription, which can be detected by the silver staining technique (Ag-staining), allowing the visualization of NORs that were transcribed in the previous interphase besides the detection of nucleoli (Goodpasture and Bloom, 1975; Roussel *et al*., 1996; Zurita *et al*., 1998). Ribosomal genes, whether active or inactive can be visualised through FISH with a 45S rDNA specific probe. In interphase, active genes appear as intranucleolar thread-like structures that emanate from perinucleolar condensed knobs that correspond to silent rRNA genes (reviewed in Neves *et al*., 2005a).

Histone post-translational modifications and chromatin remodelling factors have been clearly associated with transcriptional active rRNA genes (review in Neves *et al*., 2005b; Santoro, 2005; Huang *et al*., 2006; Preuss and Pikaard, 2007; Grummt, 2007). Epigenetic state of rRNA genes results from a complex interplay of DNA methyltransferases and histone-modifying enzymes that act in concert with chromatin remodelling complexes and RNA-guided mechanisms to define the transcriptional state of rDNA (reviewed in McStay and Grummt, 2008).

In all eukaryotic cells, the arrays of 45S rRNA units are arranged in tandem with head-to tail orientation, with an intergenic spacer (IGS) separating the ribosomal rDNA cistron (18S-ITS1-5.8S-ITS2-25S). This coordinated orientation is such that the direction of transcription was found to be oriented toward the centromere in the chromosomes 2 and 4 of *A. thaliana* (Copenhaver and Pikard, 1996 a; Lin *et al*., 1999;

Mayer *et al*., 1999). Also the terminal NOR of chromosome 9 of *Oryza sativa* have their telomere-proximal rRNA genes oriented with their 5' ends nearest the chromosome ends and their 3' ends nearest the centromere (Fujisawa *et al*., 2006).

Concerted evolution has been proposed for the homogenization of rDNA sequences clusters within a species although allowing sequence divergence across species (Flavell, 1986; Gerbi, 1986). This homogenization seems to be mediated by interchromosomal recombination between NORs on different chromosomes and intrachromosomal unequal recombination and/or gene conversion (review in Charlesworth *et al*., 1994). However length polymorphisms were found in the intergenic spacers of several plant species (Oono and Sugiura, 1980; Allard *et al*., 1990; Bellarosa *et al*., 1990; Gruendler *et al*., 1991; Luschnig *et al*., 1993). Species with more than one rDNA *locus* can present differences between those *loci* in what concerns IGS length polymorphisms as in NOR2 and NOR4 of *A. thaliana* (Copenhaver and Pikard, 1996b). The IGS sequences often contain differences that allow for determining genetic relationships between plant varieties, populations, and individuals (Rogers and Bendich, 1987; Jorgansen and Cluster, 1988; Schaal and Learn, 1988).

In this chapter, we study NOR structure and major NOR activity, in Fagaceae species. Molecular characterization, sequential silver staining and FISH, and immunodetection were performed in species of the three genera of *Fagus*, *Quercus* and *Castanea* in order to correlate structure, activity and epigenetic marks in the Fagaceae NORs.

## IV.6. Materials and Methods

### IV.6.1.Plant material

Root-tips were collected mainly from germinated acorns or from seedlings in pots from the following species: *Fagus sylvatica* L., *Quercus suber* L., *Quercus ilex* L. subsp *rotundifolia* (Lam) O. Schwarz, *Quercus acutissima* Carruth., *Castanea sativa* Mill. *Castanea crenata* Sieb. & Zucc, and *Castanea mollissima* Bl.

### IV.6.2. Slide preparations

**a** – Root tips were fixed in ethanol: glacial acetic (3:1 v/v) and further transferred to fresh FAA fixative (50% ethanol - 37% formaldehyde - glacial acetic acid, 18:1:1 v/v/v) and maintained at -20ºC for 2 to 3 days before being sequentially silver stained, hybridized through FISH with a 45S rDNA probe and immunolabelled to detect methylated cytosines. Alternatively, root tips were fixed only in ethanol: glacial acetic (3:1 v/v), and further hybridized through FISH with a 45S rDNA probe and immunolabelled to detect methylated cytosines.

**b** – Simultaneous detection of H3K9me2 through immunolabelling and DNA FISH was performed after fixation in fresh 4% (w/v) formaldehyde solution.

### IV.6.3. Silver staining, DNA:DNA FISH and Immunolabelling

Silver staining was performed according to the procedure described in section **II.4**.

For DNA: DNA FISH analysis the p*Ta*71 probe corresponding to the 45S rDNA was used in all species. Probes obtained from agarose electrophoresis isolated bands corresponded to sequences in between 25S and 18S rDNA sequences and further labelled by PCR were used in cells from species which generated them. Probes labelled with digoxigenin were detected with anti-digoxigenin antibody conjugated to FITC, and biotin-labelled probes were detected with streptavidin-Cy3 as described in section **II.5**.

Immunolabelling of methylated cytosines and H3K9me2 were performed according to procedures described in section **II.6.3** and **II.6.2**, respectively.

### *IV.6.4. Cell analysis and image acquisition*

Silver staining slides were analysed using a bright field light microscope (Axioskop 2, Zeiss). Hybridized and immunodetected slides were analysed with an epifluorescence microscope (Axioskop 2, Zeiss), and images were collected using an AxioCam digital camera (Zeiss) controlled by AxioVision 3.0 and assembled using Adobe Photoshop 6.0 (Adobe Systems Inc.).

### *IV.6.5. Amplification and cloning of sequences*

Amplification of sequences from genomic DNA of *Castanea sativa*, *Quercus ilex* subsp. *rotundifolia, Quercus suber* and *Fagus sylvatica* with primers designed for the amplification of the IGS of the 45S rDNA unit was made according to section **II.2.1**. Cloning of sequences and isolation and purification of plasmid DNA was made according to section **II.2.3** and **II.2.4**.

## *IV.7. Results*

### *IV.7.1. Morphology of the Fagaceae major NORs*

In this work the term pseudosatellite is used to describe the nucleolar chromosome satellite region with a characteristic spherical morphology that is enriched in rDNA sequences detected by FISH with an rDNA probe. All the *Quercus* and *Castanea* species, except *C. mollissima,* show one nucleolar chromosome pair with a clear secondary constriction (SC) that ends in a notorious pseudosatellite (Fig. IV.1 a-l). In metaphase chromosomes the pseudosatellite region presents an intermediary DAPI staining (Fig. IV.1 a,c,e,i,k, white arrows) between the fainter colour of the SC and the

stronger label of the rest of the chromosome probably reflecting differences in DNA composition and chromatin organization.

*Fagus sylvatica*, although presenting all NORs in a terminal position, does not have the pseudosatellite region (Fig. IV.2, white arrows), and in *C. mollissima* major nucleolar chromosomes have a conventional satellite presenting a strong DAPI labelling and a complete absence of rDNA in this domain (Fig. IV.1 m,n).

Pseudosatellite regions show different dimensions in the species studied: pseudosatellites from *Quercus acutissima* (Fig. IV.1 c) and *Quercus suber* (Fig. IV.1 e, g) are the smallest followed by *Quercus ilex* subsp. *rotundifolia* (Fig. IV.1 a) and *Castanea crenata* (Fig. IV.1 k) while *Castanea sativa* (Fig IV.1 i) shows the largest ones.

Analysis of complete metaphase plates reveals that homologous pseudosatellites can present size homomorphism or heteromorphism (Fig. IV.3). Prometaphase or metaphase cells with homomorphic (Fig. IV.3 a, e, white arrows) or heteromorphic pseudosatellites (Fig. 3 c, g, white arrows) can be detected and the rDNA nature of these regions was confirmed with FISH (Fig. IV.3 b, d, f, h).

Prometaphase pseudosatellites can lay very far away from the rest of the chromosome (Fig. IV.1 g,h; Fig. IV.3 a,b; Fig. IV.3 c,d) what often lead to misinterpretations (Ohri and Ahuja, 1990; Zoldoš *et al*., 1998; Chen *et al*., 2007) during chromosome number and chromosome type evaluation.

**Figure IV.1** – Major NOR chromosomes of *Quercus ilex* subsp. *rotundifolia* (**a**,**b**); *Quercus acutissima* (**c**,**d**); *Quercus suber* (**e**,**f**,**g**,**h**); *Castanea sativa* (**i**,**j**); *Castanea crenata* (**k**,**l**); and *Castanea mollissima* (**m**,**n**). DAPI staining (**a, c, e, g, i, k, m**) reveal chromosomes structure with the pseudosatellite (white arrow). FISH with 45S rDNA probe (**b, d, f, h, j, l,n**) on the same chromosome reveals the rDNA nature of the pseudosatellite

**Figure IV.2** –Prometaphase chromosomes of *Fagus sylvatica*. DAPI staining (**a**) reveals the presence of chromocenters. No secondary constriction can be distinguished. FISH with 45S rDNA probe discloses the four terminal NOR *loci*, (**b** white arrows). Bars= 5 μm

**Figure IV.1**

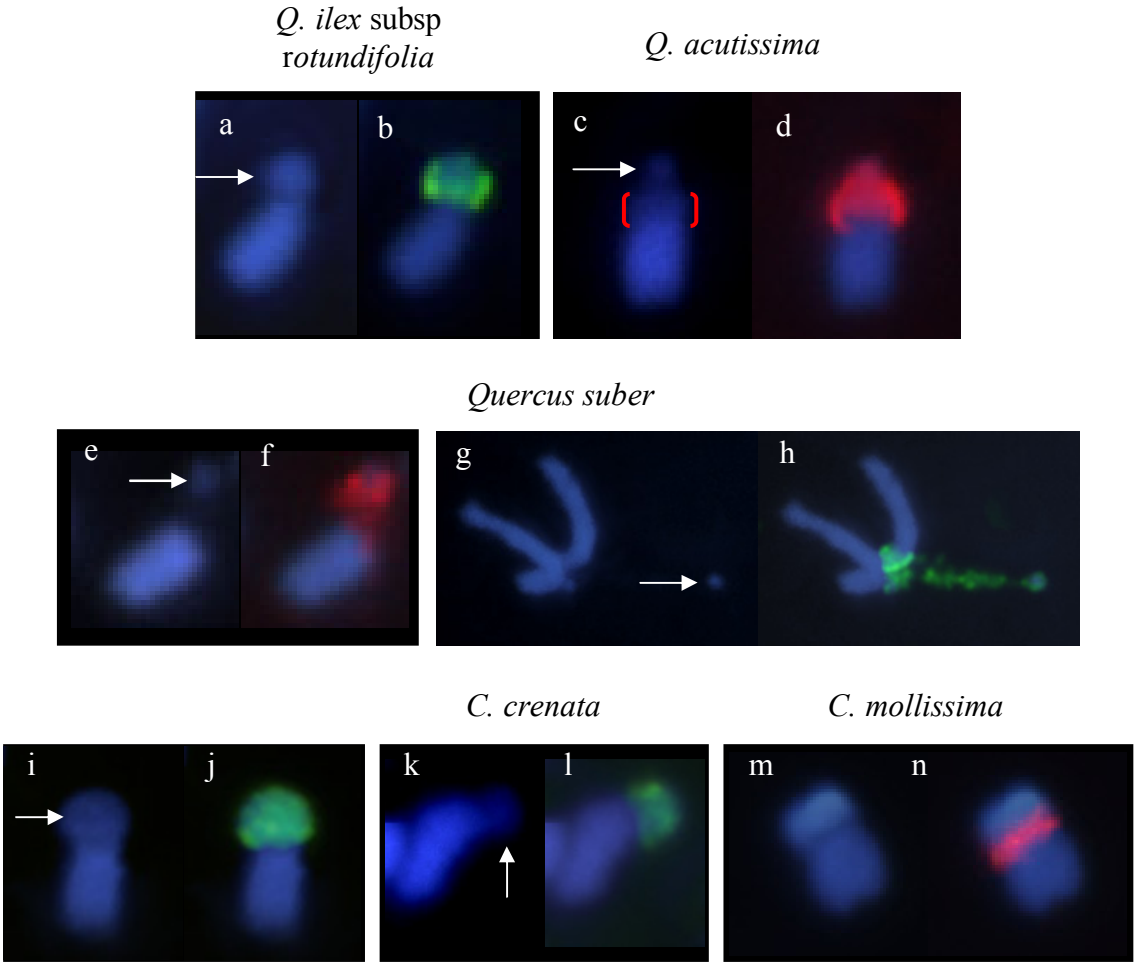



**Figure IV.2**

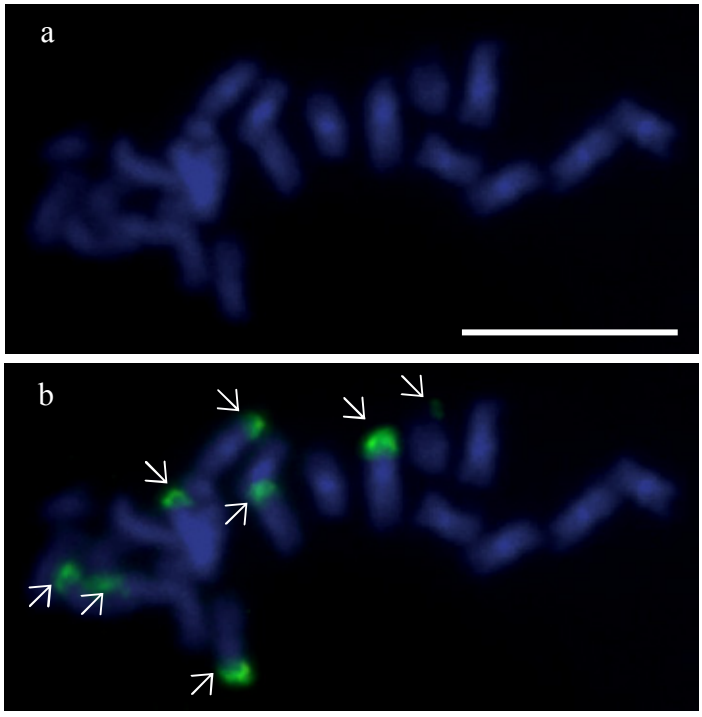
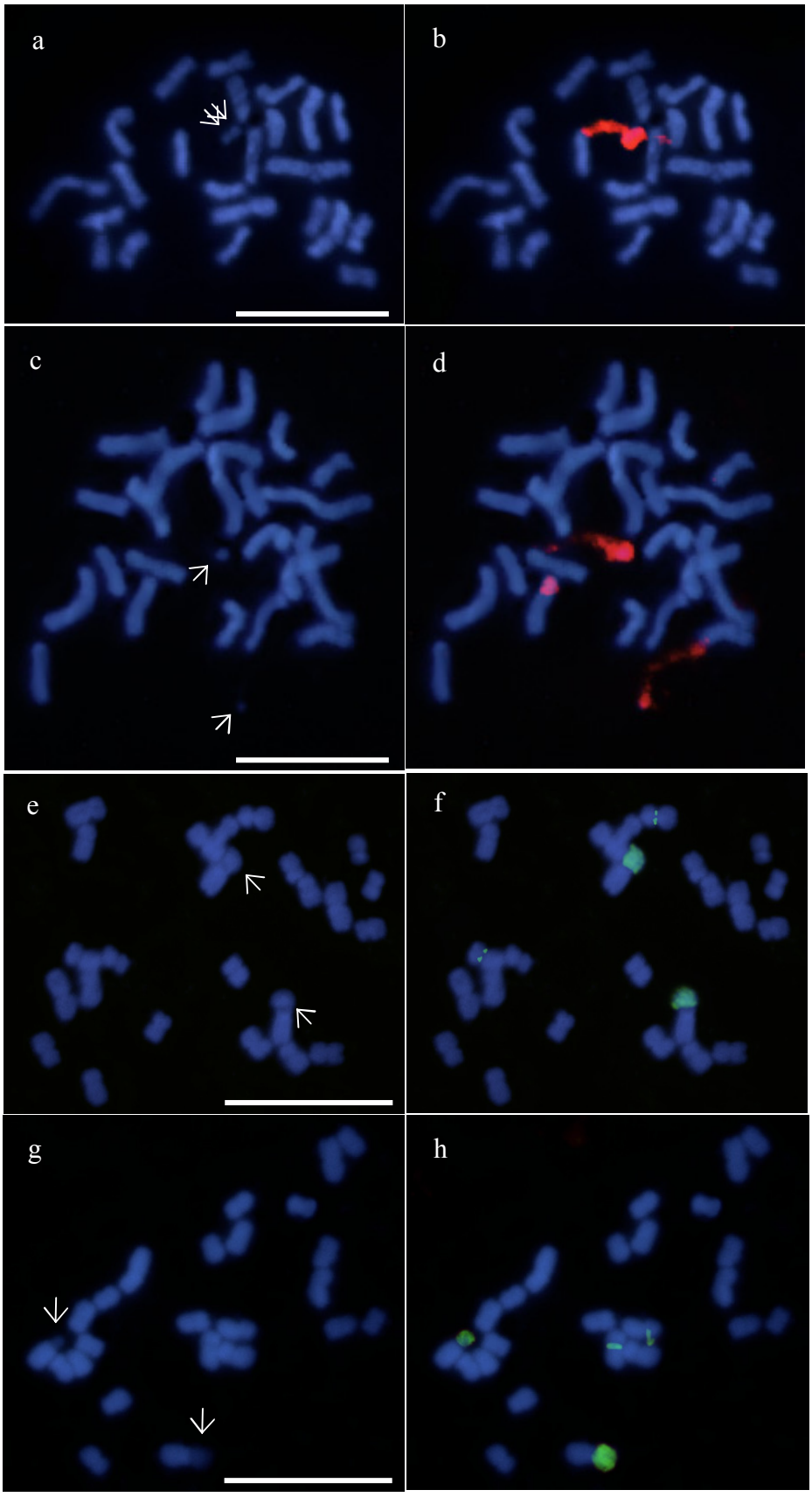


**Figure IV.3** – Prometaphase and metaphase root-tip chromosomes of Fagaceae species with small and large pseudosatellies showing size homo- or heteromorphisms detectd with 45S rDNA probe (green signal). Prometaphase of *Quercus suber* and a metaphase of *Castanea sativa* showing two pseudosatellites with similar sizes (**a,b** and **e,f**, white arrows**)**. Prometaphase of *Quercus suber* and metaphase of *Castanea sativa* presenting two pseudostellites with different sizes (**c,d** and **g,h,** white arrows**)**. DNA is counterstained with DAPI. Bars= 10 μm.

**Figure IV.3**



### IV.3.2-Chromatin organization of Fagaceae NORs

Sequential silver stain and 45S rDNA FISH on metaphase chromosomes of *C. sativa* (Fig. IV.4) revealed that only one chromosome pair exhibits the typical Ag-staining (Fig. IV.4 a, arrows), confirming that only the major NOR is active. FISH on these chromosomes shows that the pseudosatellites display very strong rDNA label (Fig. IV.4 b) indicating a marked presence of these sequences in this chromosome region although no Ag-staining could be detected.

Prometaphase chromosomes are useful structures to disclose the high resolution microscopic organization of rDNA chromatin in Fagaceae (Fig. IV.5). Silver staining (Fig. IV.5 a) and rDNA FISH (Fig. IV.5 b) were used together on prometaphase chromosomes to reveal distinct NORs domains in *C. sativa*.

Four distinct domains can be detected from centromere toward telomere in major NORs of prometaphase chromosomes:

First domain: rDNA silver stained condensed region at proximal-centromere end of the secondary constriction (Fig. IV.5 d, bright blue arrows) also detected by FISH (Fig. IV.5 e, bright blue arrows);

Second domain: decondensed region slightly silver stained (Fig. IV.5 d, greenish arrows) but clearly detected by FISH (green signal in Fig IV.5 e, greenish arrows);

Third domain: rDNA silver stained condensed region at the distal-centromere end of the secondary constriction in the pseudosatellite (Fig. IV.5 d, black arrows);

Fourth domain: the terminal chromosome region strongly stained trough FISH (green label in Fig. IV.5 e, orange arrows) and silver stained lighter than the rest of the chromosome corresponding to the pseudosatellite (Fig. IV.5 d, orange arrows).

Minor NORs (Fig. IV.5 a-c, white arrows and inset) stay far from the large prometaphasic nucleolus.

Interphase cells of *Quercus suber*, were silver stained in order to detect the maximum number of nucleoli per cell, and therefore infer about the number of active NORs. Diploid plants show the maximum number of two nucleoli (Fig. IV.6 a) while triploid cells present a maximum of three (Fig. IV.6 a,b), each one presumably corresponding to the nucleolar activity of a major NOR. Triploid cells with three nucleoli present marked nucleoli heteromorphism with two large nucleoli and a third small nucleolus.

45S rDNA FISH in prophase chromosomes of *Quercus suber* disclose NOR domains consistent with the first, second and fourth domains described for *C. sativa* (Fig. IV.7), namely the condensed region at proximal-centromere end of the secondary constriction (first domain, bright blue arrows); a decondensed region (second domain, greenish arrows); and the terminal chromosome region corresponding to the pseudosatellite (fourth domain, orange arrows). Prophase nuclei moreover reveal two separated rDNA strings with several foci emerging from the condensed first domain at the periphery of the nucleolus (Fig. IV.8). These two strings correspond to sister chromatids on the intranucleolar decondensed domain of the secondary constriction.

Pseudosatellite and the first domain (the proximal centromere end of the SC) can markedly vary in size between species being particularly evident in the Asiatic species *Q. acutissima* (Fig. IV.1 c, red brackets) where this domain is larger than in the other species studied.

In interphase nuclei perinuclear heterochromatic knobs can often be disclose by DAPI staining (Fig. IV.9 a-d). rDNA FISH analysis in interphase nuclei of *Q. suber* (Fig. IV.9 a,b) and *C. sativa* (Fig. IV.9 c,d) revealed one to four perinucleolar knobs (corresponding to first and fourth domains). The two major perinucleolar knobs corresponding to the pseudosatellites can appear individualized (Fig. IV.9 a,b) or fused (Fig. IV.9 c,d). The fusion of the pseudosatellites can also be observed in prophase chromosomes (Fig. IV.9 e,f). When pseudosatellites are fused, the third NOR domain of each chromosome can stay individualized (Fig. IV.9 d,e, arrows) or can be joined (Fig. IV.9 f, arrow). Also homologous pseudosatellites can fuse in opposite directions (Fig. IV.9 e, arrows) or the third NOR domains can overlap (Fig. IV.9 f, arrow).

**Figure IV.4** – Metaphase root-tip chromosomes of *Castanea sativa.* Silver staining, labelling the secondary constriction of major NORs (**a**, black arrows) and sequential FISH with 45S rDNA and 5S rDNA probes (**b,** green and red signal, respectively). DNA is counterstained with DAPI. Bar= 10 μm

**Figure IV.5** – Prometaphase root-tip chromosomes of *Castanea sativa.* Sequential silver staining **a)**, FISH with 45S rDNA probe **b)**, and immunodetection of methylated cytosines **c)**. Minor NORs are visible (**a-c**, white arrows, and inset). The large nucleolus appears light brown **a)**. A detail is shown in **d)**,**e)**,**f)** where distinct NORs domains can be seen: condensed centromere-proximal domain indicated by bright blue arrows), decondensed domain indicated by greenish arrows); pseudosatellite condensed domain in the distal region of the secondary constriction indicated by black arrows); Pseudosatellite indicated by orange arrows). DNA is counterstained with DAPI (**b)** and **e)**. Bar=10 μm

**Figure IV.4**

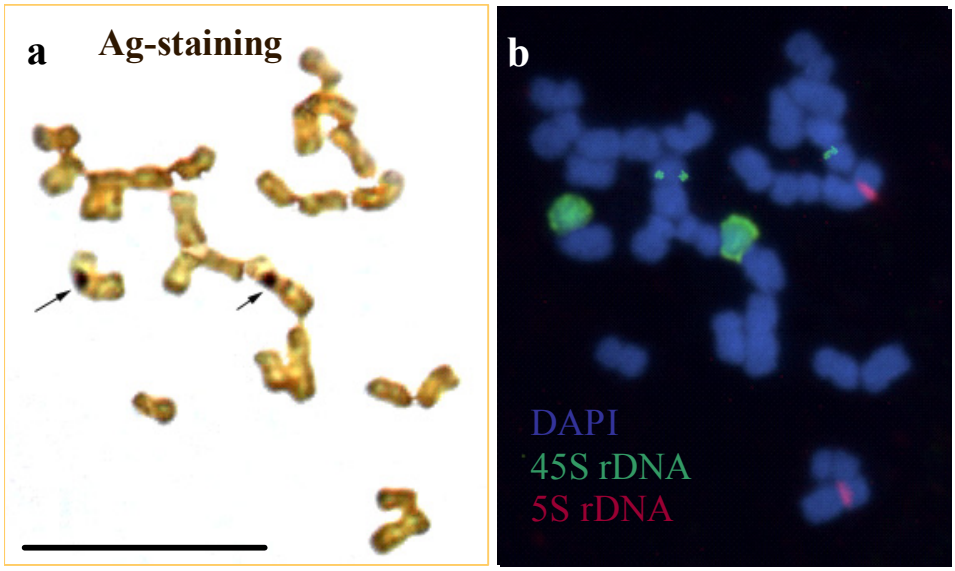



**Figure IV.5**

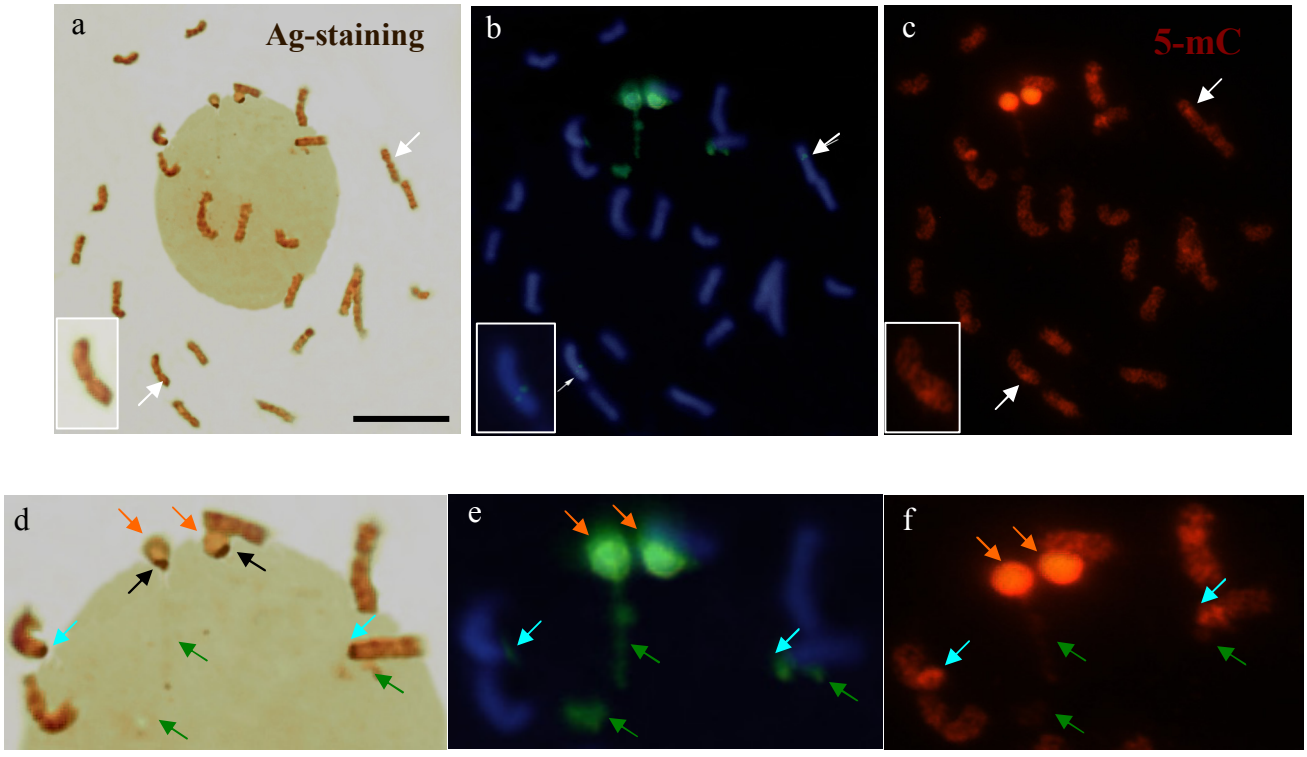


**Figure IV.6** –Diploid **a)** and triploid **b)** meristematic interphase nuclei of *Quercus suber*. Silver staining reveals two equal size **a)** and three different size **b)** nucleoli. FISH with 45S rDNA probe **c)** reveal foci inside all nucleoli which appear unstained with DAPI. Bar=10 μm.

**Figure IV.7**- Prophase meristematic root-tip chromosomes of *Quercus suber*. FISH with 45S rDNA probe (green label) reveal the NOR major chromatin domains: first domain (bright blue arrows); second domain (greenish arrows); fourth domain (orange arrows). DNA is counterstained with DAPI. Bar= 10 μm

**Figure IV.8-** Prophase meristematic root-tip chromosomes of *Quercus suber*. FISH with 45S rDNA probe shows two separated sister chromatids in the intranucleolar decondensed domain. Bar= 10 μm

**Figure IV.6**

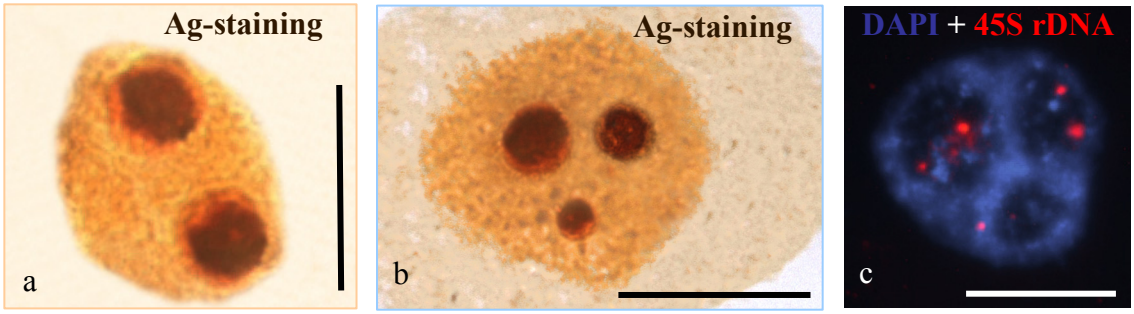



**Figure IV.7**

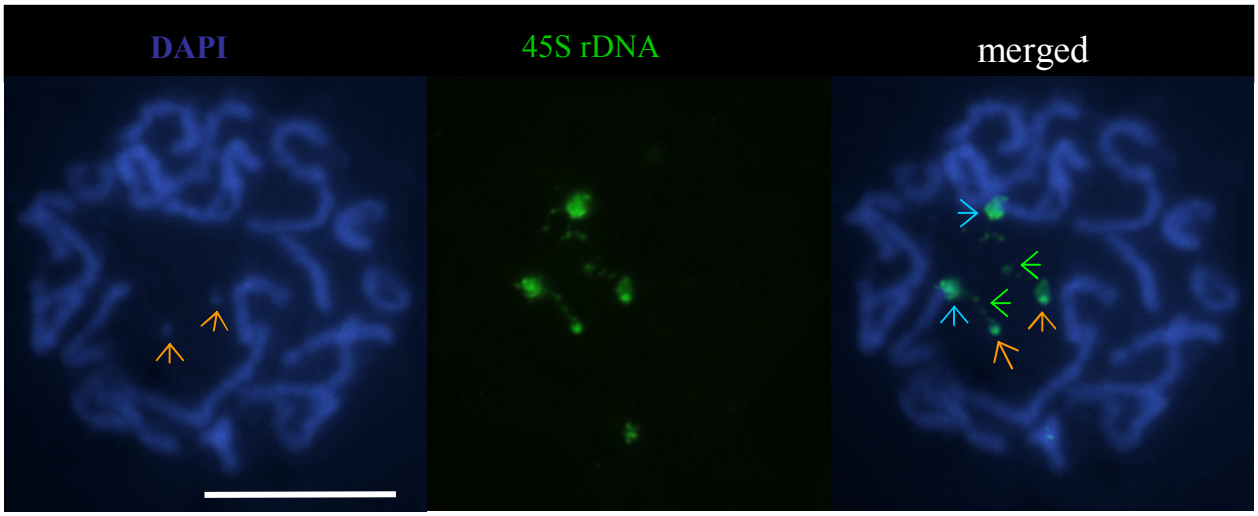



**Figure IV.8**

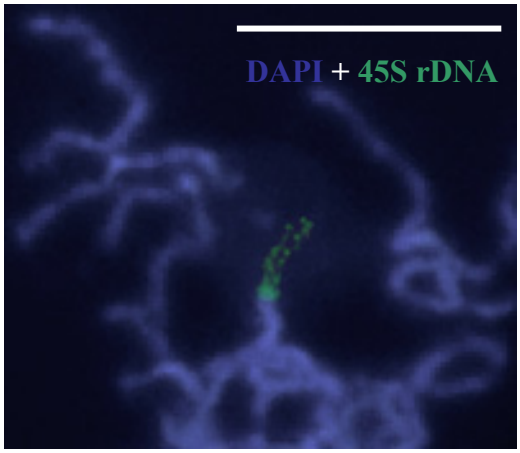
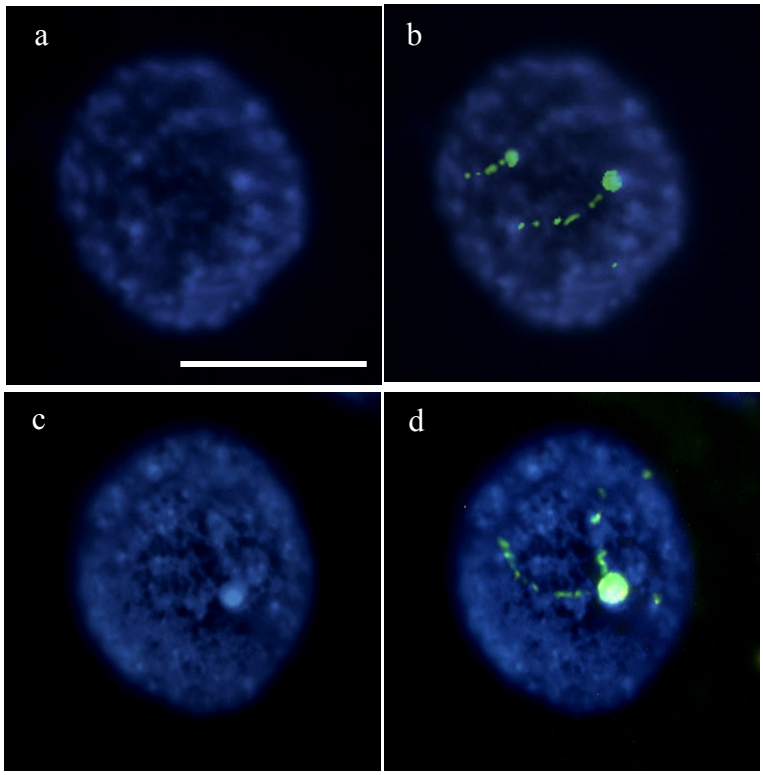
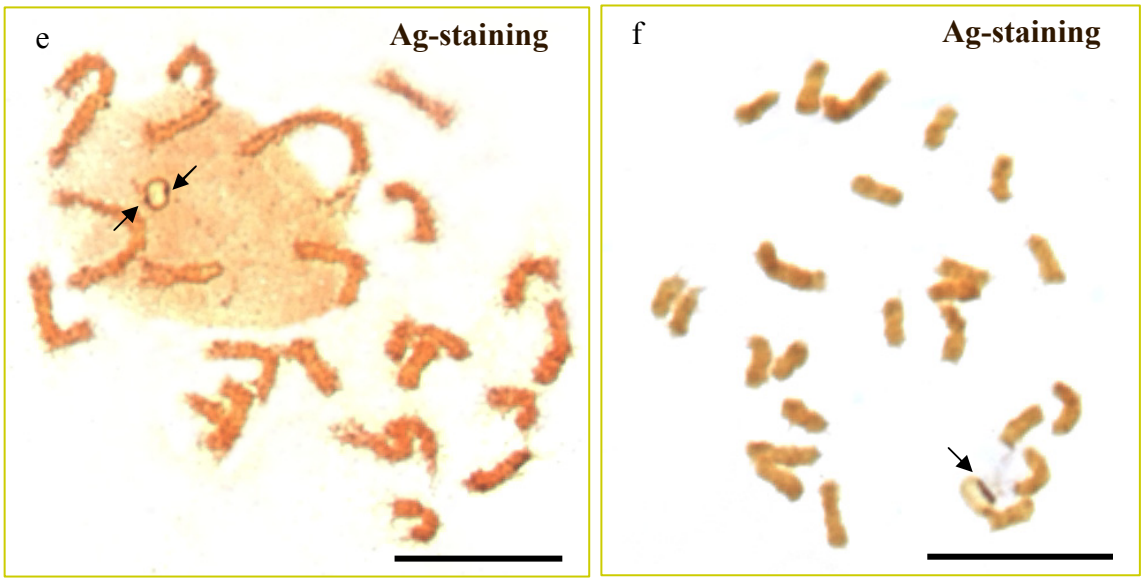


**Figure IV.9** – Interphase meristematic root-tip nuclei of *Quercus suber* **(a,b)** and *C. sativa* **(c,d)**, and prophase root-tip chromosomes of *C. sativa* (**e**,**f**). DAPI staining allows for detection of separated perinucleolar knobs with intenser label and with different sizes **a)**. FISH with 45S rDNA probe (green signal) confirms the ribosomal nature of these knobs from which emanate two intranucleolar decondensed dotted strings **b)**. Perinucleolar knobs can fuse **c), d)**. Silver staining of prophase and prometaphase chromosomes allows for the detection of chromosomes orientation. Chromosomes can fuse in opposite directions **(e)**, black arrows), or can joined together in the same orientation **(f)**, black arrow). Bars=10 μm

**Figure IV.9**



### *IV.3.3-Comparative epigenetic patterns of Fagaceae NORs*

The distribution patterns of epigenetic marks associated with gene silencing were evaluated by immunodetection in several species of the Fagaceae family. In order to detect methodology interference in methylated cytosine immunolabelling patterns, several experiments were conducted where immunodetection was performed either before or after FISH. Immunolabelling patterns revealed to be more consistent when immunodetection was performed after FISH procedure; therefore the methylated cytosines detection in all species was performed according with this order.

In *Fagus sylvatica* the major heterochromatic interphase blocks enriched in 5mC (Fig. IV.10 A) and H3K9me2 (Fig. IV.10 B) co-localised to six out of the eight rDNA sites. One euchromatic locus is 5-mC (Fig. IV.10 A, white arrows) and H3K9me2 (Fig. IV.10 B, white arrows) negative revealing its active nature. Silver staining performed on interphase nuclei failed to reveal in all cells analysed more than one small nucleolus per nucleus (Fig. IV.11).

*Quercus suber* (Fig. IV.12) and *Quercus ilex* subsp. *rotundifolia* interphase nuclei (Fig. IV.13) reveal no stronger labelling of α-5-mC at the perinucleolar knobs (pseudosatellites) whether small but fused (Fig. IV.12) or large and individualized (Fig. IV.13). *Castanea mollissima* chromosomes have a true satellite which has an heterochromatic appearance at metaphase (Fig. IV.1 m,n), prometaphase (Fig. IV.14 a,b) or prophase (Fig. IV.14 c-f). 5-mC labelling pattern is uniform along all chromosomes, including the satellite region being absent in the secondary constriction (Fig. IV.14 b,f green colour, green arrow), and more intense in the NOR domain proximal to centromere (Fig. IV.14 f, orange colour, white arrow). Interphase nuclei of *Castanea crenata,* which also presents a notorious pseudosatellite, exhibit a more intense α-5-mC labelling in the perinucleolar knobs (Fig. IV.15 A, orange colour, white arrow) and slight increased label was also detected in the distal domain of the pseudosatellite (Fig. IV.15 B, orange colour, white arrow).

*Castanea sativa* nuclei (Fig. V.3 a,b,c) and prometaphase (Fig. IV.5 f, orange arrows) show massive methylation in the pseudosatellite region. The similar intensities of 5-mC labelling were detected in both heteromorphic pseudosatellites (Fig. IV.16). Secondary constriction failed to reveal any 5-mC label (Fig. IV.16, white arrow), although at prometaphase a very residual labelling could be detected in the decondensed string domain (Fig. IV.5 f, greenish arrows). The remaining chromosomes present a

relatively constant and fainter level of 5-mC except for the first domain (Fig. IV.5 f, bright blue arrows) and the minor NOR loci (Fig. IV.5 c, white arrows, and inset), where a high level of methylated cytosines can be detected. However, the enrichment of α-5-mC labelling in the minor NOR is less consistence, being sometimes absent in the various cells observed.

**Figure IV.10 –** Interphase root-tip nuclei of *Fagus sylvatica* with 45S rDNA probe FISH (green) followed by immunodetecion of methylated cytosines -5-mC (red) **A)**. Immunodetection of H3K9me2 (red) followed by FISH with 45S rDNA probe (green) **B)**. Most of the rDNA colocalize with DAPI positive blocks and is enriched in 5-mC and H3K9me2 (orange). rDNA domains that lack 5-mC and H3K9me2 are also observed (white arrows). Bar=5 μm

**Figure IV.11** - Interphase meristematic root-tip cell of *Fagus sylvatica.* Silver staining reveals one small nucleolus. Bar= 5 μm

**Figure IV.12** - Interphase meristematic root-tip nucleus of *Quercus suber*. FISH with 45S rDNA probe (green) shows fusion of pseudosatélites. Immunodetection of methylated cytosines -5-mC (red) was followed after FISH. The pseudosatellite domain shows no enrichment of 5-mC (green detected in the superimposed 5-mC and 45S rDNA). Bar=10 μm

**Figure IV.13 -** Interphase meristematic root-tip nucleus of *Quercus ilex* subsp. *rotundifolia.* FISH with 45S rDNA probe (green) shows four signals co-localized with the heterochromatic knobs (DAPI positive) at the periphery of nucleolus; The pseudosatellite domain shows no enrichment of 5-mC (green detected in the superimposed 5-mC and 45S rDNA). Bar=10 μm.

**Figure IV.10**

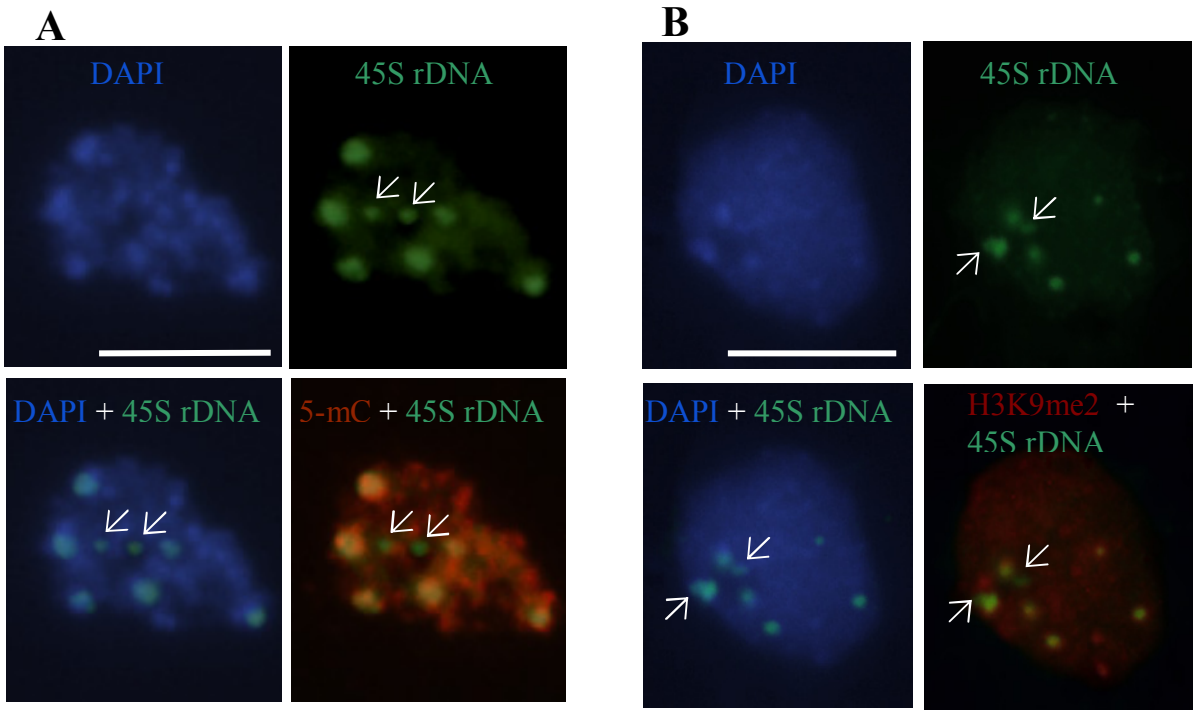



**Figure IV.11**

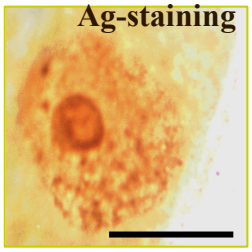



**Figure IV.12**

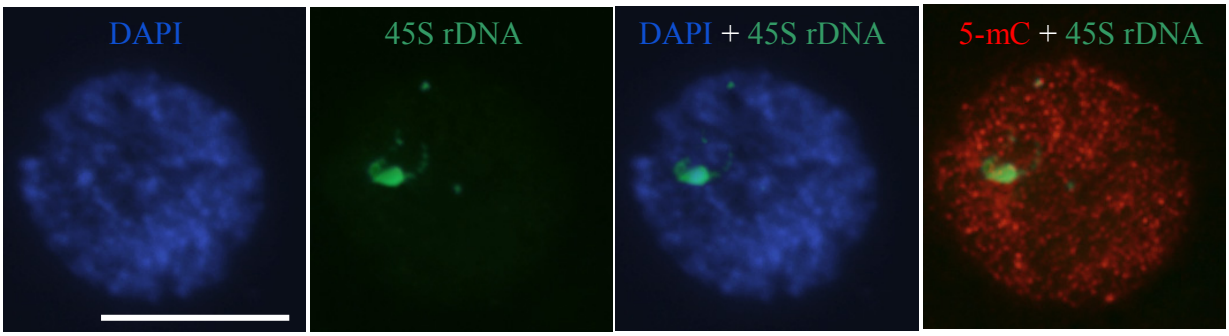



**Figure IV.13**

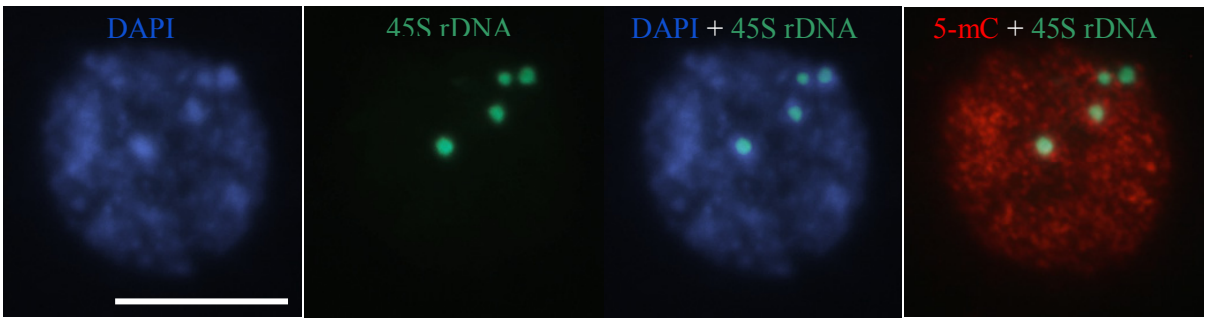


**Figure IV.14 –** Prometaphase meristematic root-tip chromosomes of *Castanea mollissima* showing two strongly heterochromatic perinucleolar satellites with DAPI **a),** arrows. Superimposed FISH with 45S rDNA probe (green) and immunodetection of methylated cytosines (red) **b)**. The 45S rDNA signal is confined to the secondary constrictions and deprived of α-5-mC labelling, which co-localizes only at the limits of the satellites but especially at the proximal centromere end (**f)** orange colour in 5-mC+45S rDNA (white arrow) at the periphery of the nucleolus (dashed circle). Bar=10 μm

**Figure IV.15 –** Interphase nucleus **(A)** and metaphase meristematic root-tip chromosomes **(B)** of *Castanea crenata.* FISH with 45S rDNA probe (green) followed by immunodetection of methylated cytosines α-5-mC (red). Two green signals co-localized with heterochromatic knobs (pseudosatellites) at the periphery of nucleolus from which emanate intranucleolar decondensed signals and there is a slight enrichment at one of the rDNA signals (orange, white arrow) **(A)**. One major NOR shows a slight enrichment of 5-mC at the pseudosatellite (orange, white arrow) **(B)**. Bar=10 μm

**Figure IV.14**

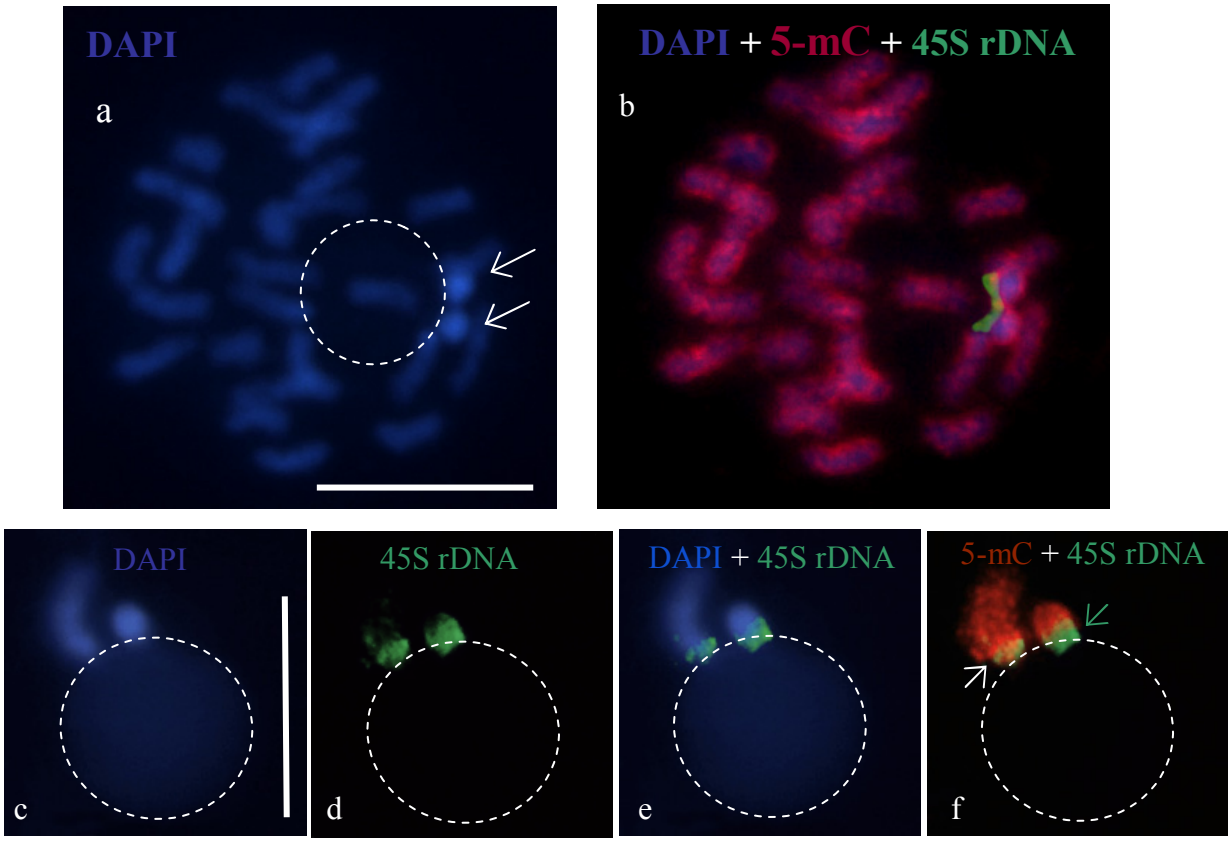



**Figure IV.15**

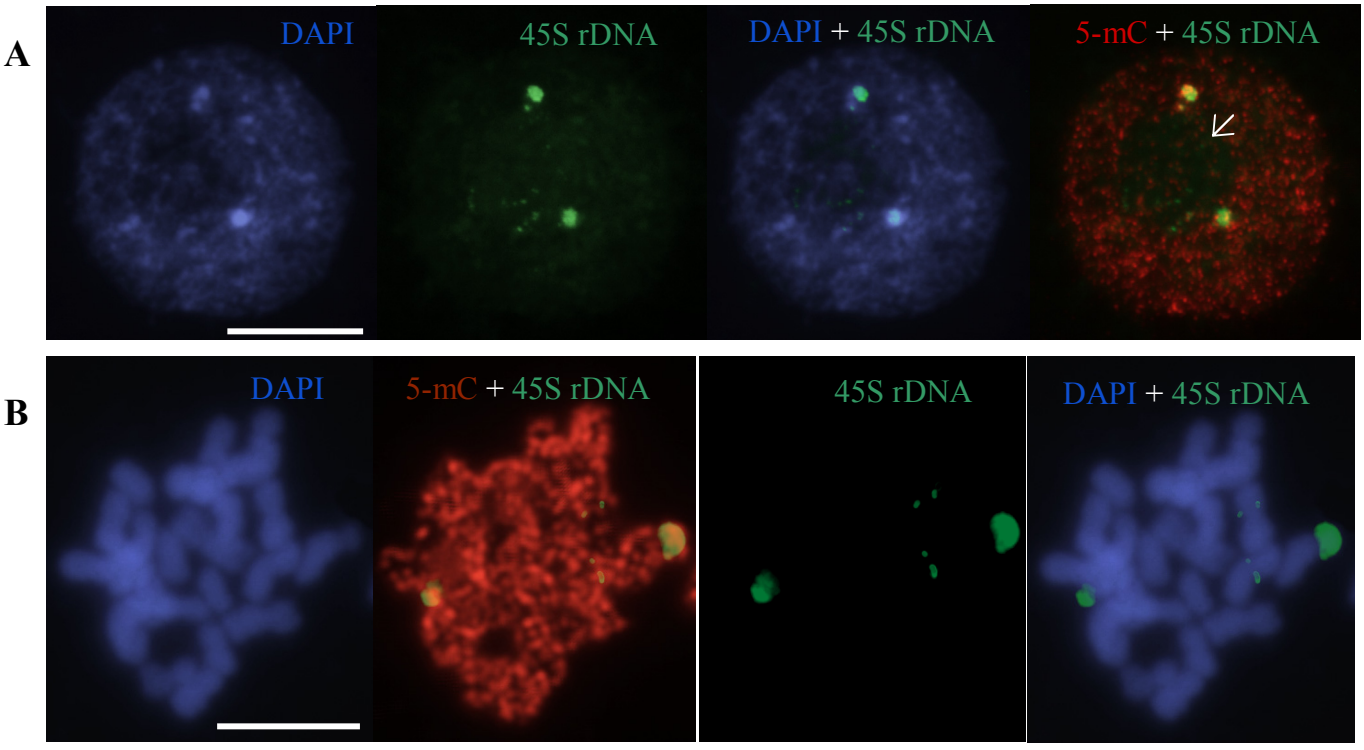
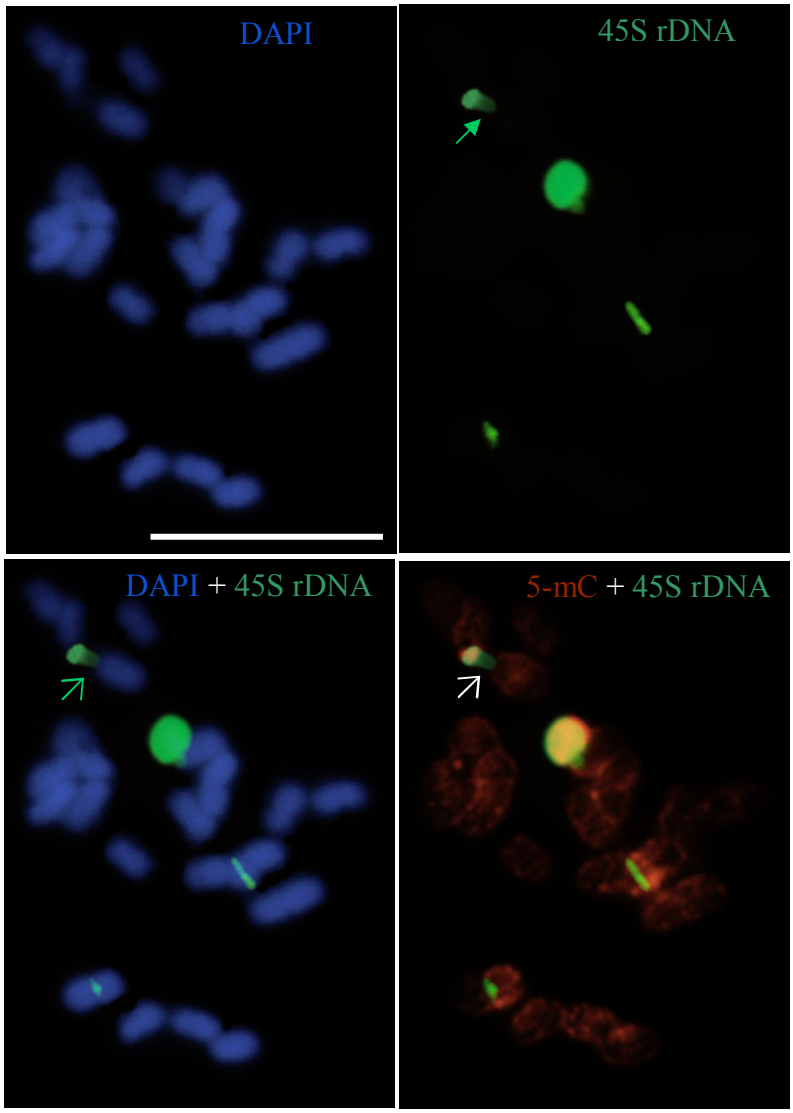


**Figure IV.16** – Incomplete meristematic root-tip metaphase of *Castanea sativa* with 45S rDNA FISH in which the secondary constriction is clearly visible (white arrow) in the chromosome bearing the smaller pseudosatellite (green arrow). The subsequent immunodetecton of methylathed cytosines shows the enrichment of this modified base both in large and small pseudosatellites. No label in the secondary constriction is detected (white arrow). Bar= 10 μm

**Figure IV.16**



### IV.7.2. Molecular characterization of Fagaceae's NORs

In order to characterize the intergenic spacers of the three Fagaceae genera, an amplification of this region was performed through PCR on genomic DNA using primers designed for the IGS flanking region: 25S forward and 18S reverse (Fig. IV.17). The *Arabidopsis thaliana* rDNA sequence unit was used to design the primers that include some nucleotides of 25S and 18S gene sequences. Amplification of genomic DNA of *Castanea sativa*, *Quercus ilex* subsp. *rotundifolia*, *Quercus suber* and *Fagus sylvatica* generate products smaller than expected (~ 2-5 kb), three products with sizes ranging from 300 to 500 bp for *Castanea* and *Quercus* species and one product around 300 bp for *F. sylvatica* (Fig. IV.18).

Fragments of *C. sativa*, *Q. ilex* subsp. *rotundifolia* and *F. sylvatica* were cloned and sequenced. Sequencing of cloned amplicons revealed that the sizes of the bigger fragments are 476 and 470bp in *C. sativa* and *Q. ilex* subsp. *rotundifolia*, respectively. These fragments are flanked by palindrome repeats of the 18S primer (Fig. IV.19 a,b) what was confirmed by PCR genomic amplification using only the 18S primer. A search of homology through the MegaBlastn algorithm revealed that these DNA sequences corresponded to the ETS region upstream the 5'end of the 18S (Fig. IV.19 scheme a,b). These sequences show around 90% of homology with extremely low e-values with the ETS regions of several *Quercus* spp. (Table IV.1). However the sequences that we obtained do not contain the complete ETS sequence, but only a portion which is truncated at 5'end by the inverted and complementary sequence of the 18S primer sequence (highlighted in dark grey). At the 3' end there is the canonical 5' end of the 18S sequence including the primer (highlighted in dark grey), nucleotides making part of the end of 18S (underlined) and a conserved sequence in plants (highlighted in lighter grey), which we believe to also make part of the end of the 18S sequence, although most databases consider it as already being part of the ETS (Fig. IV.19).

*Fagus sylvatica* single fragment although having only 305 bp belongs to the same genomic region (Fig. IV.19 c) and presents the same organization. The 18S palindromic repeat in this sequence was also confirmed by PCR using the single 18S primer. The search of homology revealed 75 to 74% of homology with the same ETS

region of *Q. robur* (Table IV.1) what agrees with the relative phylogenetic position of the three species studied.

Comparisons between the fragments of the three species with the Kalign algorithm (Fig. IV.20) revealed that *Quercus* and *Castanea* fragments display the 18S palindromic insertion in the same region of the intergenic sequence while *F. sylvatica* presents the palindromic 18S insertion nearer the 5' end of 18S gene. A ClustalW multialign of the intergenic sequences after removing the 18S inverted insertion, revealed that *C. sativa* and *Q. ilex* subsp. *rotundifolia* share 91% of homology, while each of these species with *F. sylvatica* share only 74% (Fig. IV.21).

The middle size fragment of *Castanea sativa* (351 bp) presents 317 bp of a 26S plant mitochondrial rDNA inserted in an inverted manner and flanked by the 25S and 18S rDNA primers sequences (Fig. IV.22). The BLASTn analysis show 99% of homology to several plant mitochondrial 26S rDNA sequences, including to the 26S ribosomal RNA region in the complete genome of *Arabidopsis thaliana* mitochondria (NC_001284.2..GI: 26556996) with an $E=2^{e-163}$.

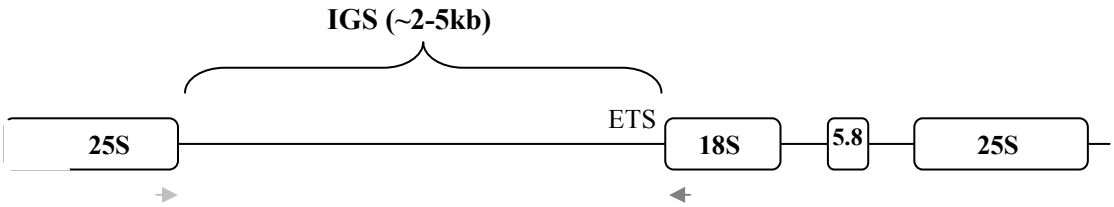
The smaller fragment amplified in *Quercus* spp. and *C. sativa* (Fig. IV.18) was not sequenced. Each of the three fragments amplified in *C. sativa* and the single *F. sylvatica* fragment were isolated from the gel and used as a template to generate individual FISH probes through PCR, and were further hybridized in *C. sativa* and *F. sylvatica* meristematic root-tip cells. In *C. sativa* both big and small PCR fragments produce similar hybridization patterns (Fig. IV.23 A,B): as expected these probes hybridized with both major and minor NORs (Fig. IV.23 A,B,a) revealing the same pattern as the complete 45S rDNA probe, previously used in Figs. IV.3 f,h; IV.4 b; IV.5 e; IV.9 d and IV.16 ; labelling all domains of the active NOR, including the decondensed intranucleolar domain (Fig. IV.23 A,b, B,b). Hybridization with the mitochondrial-like fragment in *C. sativa* interphase nuclei is restricted to the condensed perinucleolar knobs (Fig. IV.23 C), and no signal was detected in the other NOR domains.

FISH with the probe produced from *F. sylvatica* genomic DNA in *F. sylvatica* meristematic root-tip nuclei revealed a pattern similar to the 45S rDNA probe, labelling all the NOR *loci* present in this species (Fig. IV.24).

**Figure IV.17** –Diagram of the 45S rDNA unit including primers annealing regions. The forward primer - Arab 25SFw (light grey) and the reverse - Arab 18SRev (dark grey) were used to amplify the intergenic spacer (IGS=NTS+ETS) between the 25S rRNA and 18S rRNA genes.

**Figure IV.18** – Agarose gel electrophoresis of PCR products using the Arab 25SFw and Arab 18SRev primers on genomic DNA of *Castanea sativa* (lane 1), *Quercus ilex* subsp. *rotundifolia* (lane 2), *Quercus suber* (lane 3) and *Fagus sylvatica* (lane 4). Molecular weight marker (lane 1Kb+)

**Figure IV.17**




Arab 25SFw : TGAGATTCAGCCCTTTGTCGCTAAG

Arab 18SRev:AGACAAGCATATGACTACTGGCAGG

**Figure IV.18**

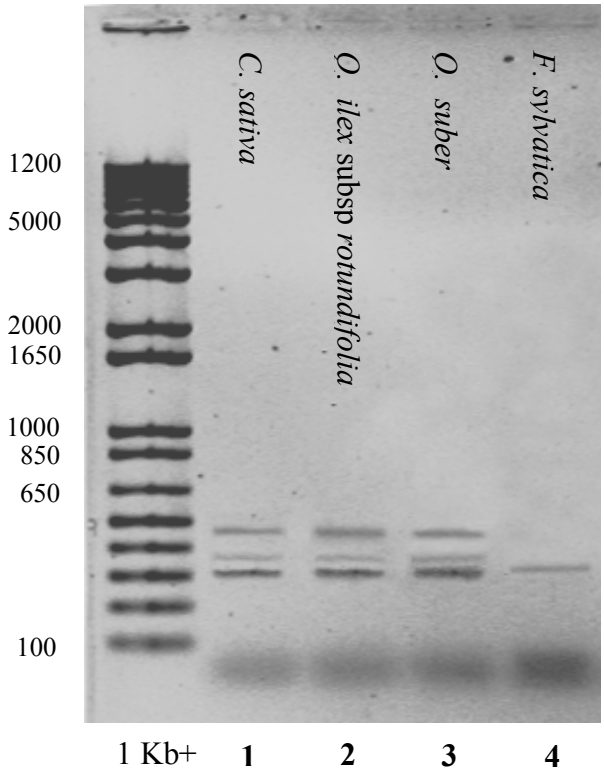
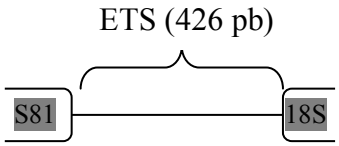


**Figure IV.19 –** Genomic DNA sequence of the 476 bp amplicon of *C. sativa* **a)** 470 bp in *Q. ilex* subsp. *rotundifolia* **b)** 305 bp in *F. sylvatica* **c).** Primers Arab18S highlighted in dark grey, the 5' end 18S sequence underlined and the conserved sequence in plants highlighted in light grey. Below each DNA sequence, a diagram showing the IGS PCR amplification restricted to a region of the ETS region -426 bp in *C. sativa* **a)**, 420 bp in *Q. ilex* subsp. *rotundifolia* **b)**, 255 bp in *F. sylvatica* **c)** which is truncated by the palindromic 18S primer sequence.

**Figure IV.19**

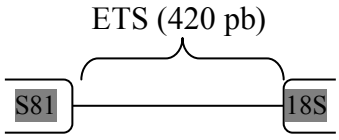
**a)**

AGACAAGCATATGACTACTGGCAGGCTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTCCC
AATCAGCGCTGTTTCGAGCGTCGCTCGACGCCTCGAACGCAATTCGGGTCCCTGTGTTGCAT
ACCTGCCTCGAAGGCACTCGTCCCTCTAGTTGATTCGTTCCTAGTCGATGCTCCTTGCGGGGT
GTCGGCAGGACCTTGAAGCCGTCCTCGCGTCCCACGCGTGCCTCGCCAAGCGTTGCCGCTGT
GGACCGCGCGGGCGTGCTCGTGGCCTCGGATGCAGAACATTATGTGGGTTTGGGGCCTCTGG
CCCCCTTTGCCAACGTACCAAGCGAGCGTCATCGCTCTGCCCCGCACGATCGCCGCGCTCGT
TTGCGCCCTTCCTTGCCTTCGGGCGAGCCTGGGCCTCCGGGCGACGCCGGCATCGACGAGGA
ATGCTACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT




**b)**

AGACAAGCATATGACTACTGGCAGGCTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTCCC
AATCAGCGTTGTTTCGAGCGTCGCTTGGACGCAATTCGGGTCCCTGTGCTGCATACCTGCCTT
GAAGGCACTCGTCCCCCTAGTTGGTTAGTTCCTAGTCGACGCTCCTCGGGGCGTCGGCAGGA
CCTAGAAGCTGTCCTCGTGTCCCACACGTGCCTTGCGGCCTTCGCGTTGCCGATGTGGACCA
CGTGGGCATGCTCGTGGCCTCGGATGCAGAACATCATGTGGGTTTGGGGCCTTTGGCCCCTT
TTGCCAACGTACCAAGCGAGCATCATCGCTTTGCCCCGCACGATCGTCGTGCTCGTCCGCGC
CCTTCCTTGCCTTCAGGTGAGCCAGGGCCTCCGGGCAGCGCCGGCATCGACGAGGAATGCTA
CCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT




**c)**

AGACAAGCATATGACTACTGGCAGGACCATGAAGTCGTCCATGTGTCCCACGGCTGCCTCGC
TTAGATGCGTTGTTGTCGTGGACCACGTGGGCGAGCTCGTGCTCTCGGTTGCAGAACATAAT
GTCGGTTCGGGGTTATCTTTACCCTTTTCGACCCAACACAAGCTTTCTCGCTTGAAACGAACG
ATCGNCNTGCCCGTCCACGAACACGTCCGTCCCTAAGGTCGGAAAAGGGTGTCGTGCGGAG
TCGGCATCGAAGAGGAATGCTACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT

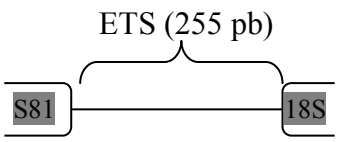


**Figure IV.20 -** Kalign multialignment of three ETS sequences isolated from *C. sativa, Q. ilex* and *F. sylvatica*

**Figure IV.20**

```
Kalign (2.0) alignment in ClustalW format


C     AGACAAGCATATGACTACTGGCAGGCTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTC
Q     AGACAAGCATATGACTACTGGCAGGCTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTC
F     AGACAAGCATATGACTACTGGCAGG----------------------------------


C     CCAATCAGCGCTGTTTCGAGCGTCGCTCGACGCCTCGAACGCAATTCGGGTCCCTGTGTT
Q     CCAATCAGCGTTGTTTCGAGCGTCG--------CTTGGACGCAATTCGGGTCCCTGTGCT
F     ------------------------------------------------------------


C     GCATACCTGCCTCGAAGGCACTCGTCCCTCTAGTTGATTCGTTCCTAGTCGATGCTCCTT
Q     GCATACCTGCCTTGAAGGCACTCGTCCCCCTAGTTGGTTAGTTCCTAGTCGACGCTCCT-
F     ------------------------------------------------------------


C     GCGGGGTGTCGGCAGGACCTTGAAGCCGTCCTCGCGTCCCACGCGTGCCTCGC----CAA
Q     -CGGGGCGTCGGCAGGACCTAGAAGCTGTCCTCGTGTCCCACACGTGCCTTGCGGCCTTC
F     ---------------ACCATGAAGTCGTCCATGTGTCCCACGGCTGCCTCGC-TTAGAT


C     GCGTTGCCGCTGTGGACCGCGCGGGCGTGCTCGTGGCCTCGGATGCAGAACATTATGTGG
Q     GCGTTGCCGATGTGGACCACGTGGGCATGCTCGTGGCCTCGGATGCAGAACATCATGTGG
F     GCGTTGTTGTCGTGGACCACGTGGGCGAGCTCGTGCTCTCGGTTGCAGAACATAATGTCG


C     GTTTGGGGCCTCTGGCCCCCTTTGCCAACGTACCAAGCGAGCGTCATCGCTCTGCCCCGC
Q     GTTTGGGGCCTTTGGCCCCTTTTGCCAACGTACCAAGCGAGCATCATCGCTTTGCCCCGC
F     GTTCGGGGTTATCTTTACCCTTTTCGACC---CAACACAAGCTTTCTCGCT-TGAAACGA


C     ACGATCGCCGCGCTCGTTTGCGCCCTTCCTTG-CCTTCGGGCGAGCCTGGGCCTCCGGGC
Q     ACGATCGTCGTGCTCGTCCGCGCCCTTCCTTG-CCTTCAGGTGAGCCAGGGCCTCCGGGC
F     ACGATCGNCNTGCCCGTCCACGAACACGTCCGTCCCTAAGGTCGGAAAAGGGTGTCGTGC


C     GACGCCGGCATCGACGAGGAATGCTACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTC
Q     AGCGCCGGCATCGACGAGGAATGCTACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTC
F     GGAGTCGGCATCGAAGAGGAATGCTACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTC


C     T
Q     T
F     T
```

**Figure IV.21 -** ClustalW multialignment of three ETS sequences of *C. sativa, Q. ilex* subsp. *rotundifolia*, and *F. sylvatica* without the 3'primer.

**Figure IV.21**

```
CLUSTAL 2.0.8 multiple sequence alignment


C.sativa        CTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTCCCAATCAGCGCTGTTTCGAGCGTCG 60
Q.ilex          CTCCGTGCTCGCGCATCGAACTGTCCGGCGTGCTCCCAATCAGCGTTGTTTCGAGCGTCG 60
F.sylvatica     ------------------------------------------------------------


C.sativa        CTCGACGCCTCGAACGCAATTCGGGTCCCTGTGTTGCATACCTGCCTCGAAGGCACTCGT 120
Q.ilex          CT--------TGGACGCAATTCGGGTCCCTGTGCTGCATACCTGCCTTGAAGGCACTCGT 112
F.sylvatica     ------------------------------------------------------------


C.sativa        CCCTCTAGTTGATTCGTTCCTAGTCGATGCTCCTTGCGGGGTGTCGGCAGGACCTTGAAG 180
Q.ilex          CCCCCTAGTTGGTTAGTTCCTAGTCGACGCTCCT--CGGGGCGTCGGCAGGACCTAGAAG 170
F.sylvatica     ----------------------------------------------------ACCATGAAG 9
                                                                     ***  ****

C.sativa        CCGTCCTCGCGTCCCACGCGTGCCTCGCCAA----GCGTTGCCGCTGTGGACCGCGCGGG 236
Q.ilex          CTGTCCTCGTGTCCCACACGTGCCTTGCGGCCTTCGCGTTGCCGATGTGGACCACGTGGG 230
F.sylvatica     TCGTCCATGTGTCCCACGGCTGCCTCGCTTAGAT-GCGTTGTTGTCGTGGACCACGTGGG 68
                  ****  * *******   ***** **       ******  *  ******* ** ***

C.sativa        CGTGCTCGTGGCCTCGGATGCAGAACATTATGTGGGTTTGGGGCCTCTGGCCCCCTTTGC 296
Q.ilex          CATGCTCGTGGCCTCGGATGCAGAACATCATGTGGGTTTGGGGCCTTTGGCCCCTTTTGC 290
F.sylvatica     CGAGCTCGTGCTCTCGGTTGCAGAACATAATGTCGGTTCGGGGTTATCTTTACCCTTT-T 127
                *  *******  ***** ********** **** **** ****         ** ***

C.sativa        CAACGTACCAAGCGAGCGTCATCGCTCTGCCCCGCACGATCGCCGCGCTCGTTTGCGCCC 356
Q.ilex          CAACGTACCAAGCGAGCATCATCGCTTTGCCCCGCACGATCGTCGTGCTCGTCCGCGCCC 350
F.sylvatica     CGACCCAACA--CAAGCTTTCTCGCT-TGAAACGAACGATCGNCNTGCCCGTCCACGAAC 184
                * **  * **  * *** *  ***** **   ** ******* *  ** ***   **  *

C.sativa        TTCCTTG-CCTTCGGGCGAGCCTGGGCCTCCGGGCGACGCCGGCATCGACGAGGAATGCT 415
Q.ilex          TTCCTTG-CCTTCAGGTGAGCCAGGGCCTCCGGGCAGCGCCGGCATCGACGAGGAATGCT 409
F.sylvatica     ACGTCCGTCCCTAAGGTCGGAAAAGGGTGTCGTGCGGAGTCGGCATCGAAGAGGAATGCT 244
                      * ** *  **   *    **    ** **   * ********* **********

C.sativa        ACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT 451
Q.ilex          ACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT 445
F.sylvatica     ACCTGGTTGATCCTGCCAGTAGTCATATGCTTGTCT 280
                ************************************


SeqA Name          Len(nt)  SeqB Name          Len(nt)  Score
=============================================================
1    C.sativa      451      2    Q.ilex        445      91
1    C.sativa      451      3    F.sylvatica   280      74
2    Q.ilex        445      3    F.sylvatica   280      74
=============================================================
```

**Table IV.1**- Megablast results of ETS amplicons of *Quercus ilex* subsp. r*otundifolia, Castanea sativa and Fagus sylvatica.*

| | *Q. ilex* subsp. r*otundifolia* (470bp) | | | *C. sativa* (476bp) | | | *F. sylvatica* (305 bp) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Query coverage (%) | % Identity | E value | Query coverage (%) | % Identity | E value | Query coverage (%) | % Identity | E value |
| *Quercus robur* 25S-18S rRNA intergenic spacer - EF208967 | 96 | 93 | 0 | 96 | 91 | $4^{e-173}$ | 93 | 75 | $2^{e-24}$ |
| *Quercus robur* 25S-18S rRNA intergenic spacer - EF208969 | 96 | 93 | 0 | 96 | 91 | $4^{e-173}$ | 93 | 74 | $5^{e-21}$ |
| *Quercus robur* 25S-18S rRNA intergenic spacer - EU555521 | 91 | 92 | $2^{e-176}$ | 92 | 89 | $1^{e-153}$ | | | |
| *Quercus petraea* 25S-18S rRNA intergenic spacer - EU555523 | 91 | 91 | $2^{e-175}$ | 92 | 90 | $1^{e-158}$ | | | |
| *Quercus petraea* 25S-18S rRNA intergenic spacer - EU555532 | 92 | 91 | $8^{e-175}$ | 92 | 90 | $1^{e-158}$ | | | |
| *Quercus petraea* 25S-18S rRNA intergenic spacer - EU555526 | 92 | 91 | $8^{e-170}$ | 92 | 89 | $2^{e-151}$ | | | |

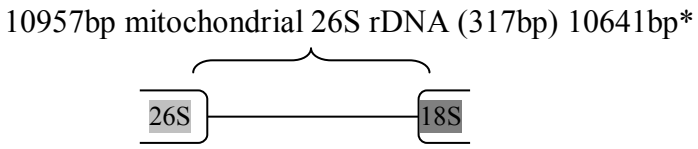
**Figure IV.22 –** Genomic DNA Sequence of the 351 bp amplicon of *C. sativa* obtained with primers Arab25SFw and Arab18SRev (light and dark grey, respectively), comprising a partial sequence (3' → 5') of the mitochondrial 26S rDNA, which shares some homology (underlined) with the primers sequences (5'→3') **a).** A diagram showing the inverted mitochondrial region and its location in the mitochondria genome of *A. thaliana* truncated by the Arab25SFw and Arab18SRev genomic primers **b)**

**Figure IV.22**

**a)**

TGAGATTCAGCCCTTTGTCGCTAAGGTCCCTAAGCAATCACTTAGTGGAAAA
GGAAGTGATCGAGCGATGACAACCAGGAGGTGGGCTTGGAAGCAGCCATC
CTTTGAAGAAAGCGTAATAGCTCACTGGTCTAGCTCCATGGCACCGAAAAT
GTATCAGGGCTCAAGTGATTCACCGAAGCGACGAGACCTTGAAAGCTGCTT
TTTCAAGTGTCAGTAGCGGAACGTTCTGTCAATCGGGGAAGGTTTTTGGTGA
CAACACCTGGAGATATCAGAAGTGAGAATGCTGACATGAGTAACGAGAAAT
CCTGTGAAAAACACGATCGCCTGCCAGTAGTCATATGCTTGTCT

**b)**




* bps from *A. thaliana* mitochondrial genome

**Figure IV.23 –** Metaphase chromosomes **a)** and interphase **b)** meristematic root-tip nucleus of *C. sativa.* FISH with ETS probes (red): 476 bp **A)**, ~300 bp **B)**, and 351 bp **C)**. The 476 bp and ~300 bp probes label both the majors and minors NORs **A,a**, **B,a**. These sequences are present in the condensed perinucleolar knobs as well as in the intranucleolar decondensed domains **A,b**, **B,b**. The 351 bp probe is restricted to the condensed perinucleolar knobs **C)**. Bars=10 μm

**Figure IV.24 –** Interphase meristematic root-tip nucleus of *Fagus sylvatica.* FISH with the ETS 305 bp probe (red) and 45S rDNA (green) reveal fusion of rDNA *loci*. Superimposed labels show a complete co-localization of both probes. Bar= 5 μm

**Figure IV.23**

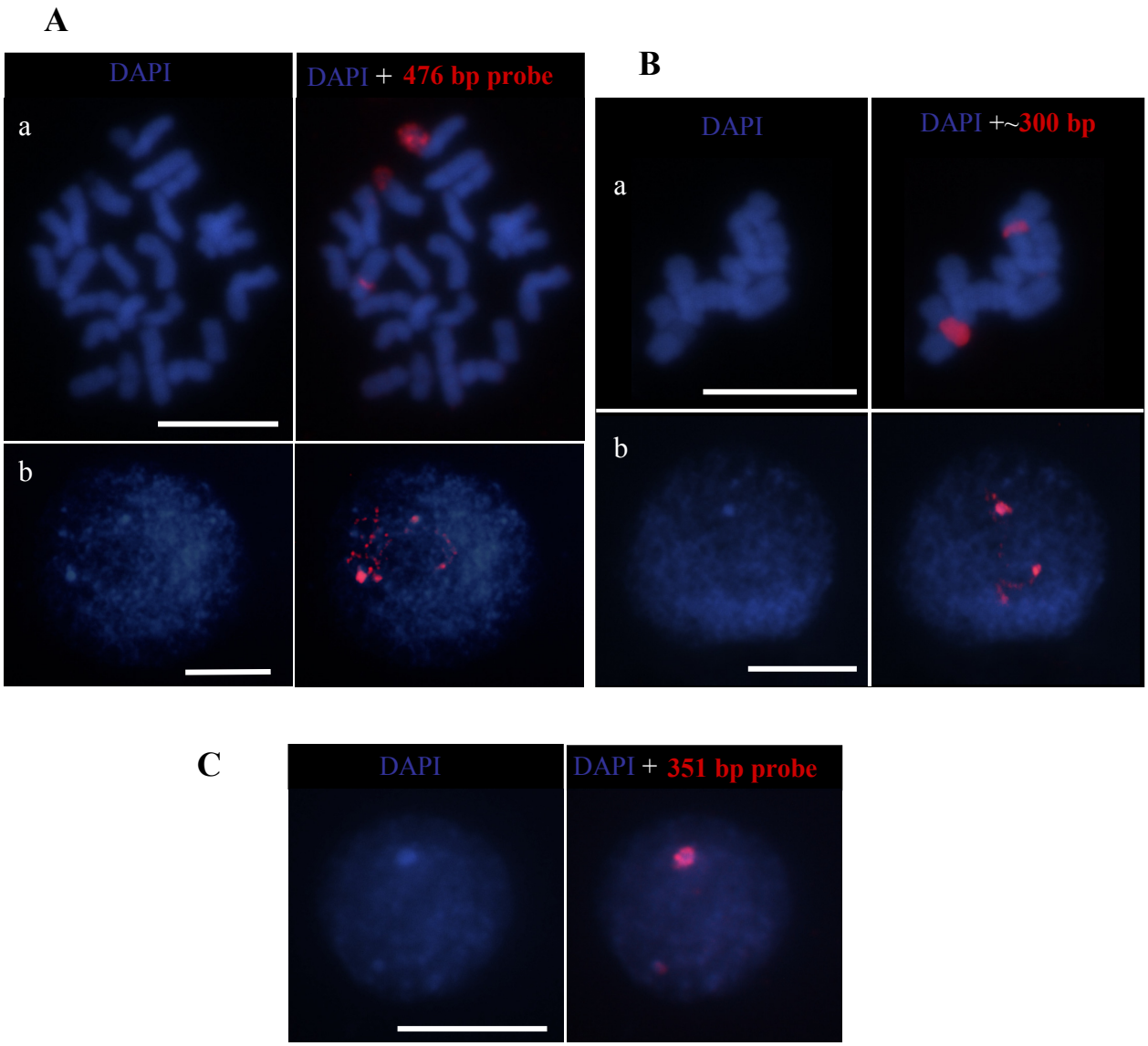
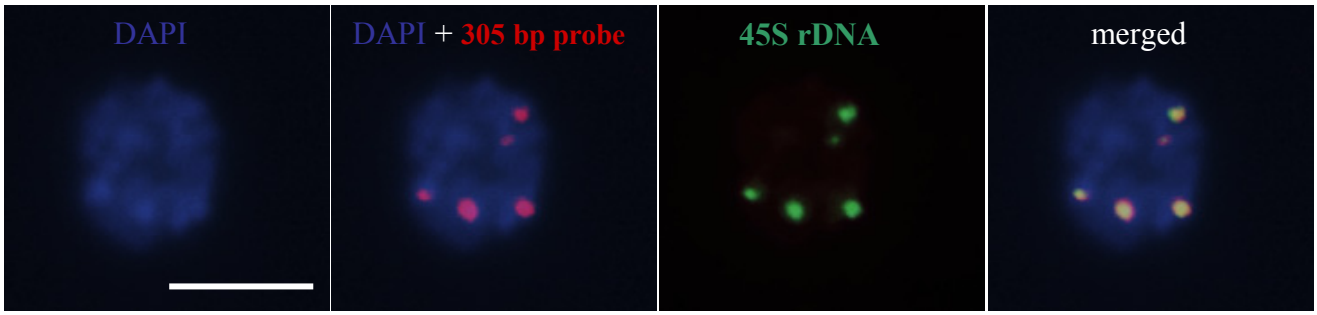



**Figure IV.24**



## IV.8. Discussion

The rRNA gene activity was evaluated by silver staining in all the Fagaceae species studied. *Castanea* species as well as diploid oaks present cells with one or two nucleoli, which are in agreement with the nucleolar activity previously reported for *Quercus robur* (Besendorfer *et al*., 1995), while the triploid *Q. suber* revealed cells with three nucleoli indicating that all the major NORs are active. In these cells a clear nucleolar heteromorphism can be detected with two large nucleoli and a smaller one reflecting differential expression of rRNA genes in homologous chromosomes (Caperta *et al*., 2002). Although presenting the double loci of 45S rDNA, in *Fagus sylvatica* only one small nucleolus could be observed. Nucleoli fusion is a common process which was only recently understood in human cells (Tsai *et al*., 2008). According to this study the two ends of the nucleolar scaffold protein hNopp140 are likely folded to a proximal position to create an anchoring domain. As the anchoring domains from many hNopp140 molecules attach to rDNA containing chromatin and acrocentric α-satellite arrays in humans, their central repeated regions are able to interact with each other, resulting in large fused nucleolus. Only a small part of the vast amount of rDNA sequences present in the NOR domains are active and the physical organization of active and inactive rDNA sequences is well known in different plant species. In rye chromosomes a centromere-proximal condensed NOR region is detected by FISH, and the NOR region decondenses progressively towards the telomeric region (Caperta *et al*., 2002), while in wheat two condensed rDNA blocks can be seen flanking the secondary constrictions of some NORs (Mukai *et al*., 1991; Morais-Cecílio *et al*., 2000), being the same pattern observed in interphase. In *Quercus* and *Castanea*, with the exception of *Castanea mollissima*, this last pattern comes to the extreme especially concerning the more distally block flanking the secondary constriction which extends to the end of the chromosome (the pseudosatellite). However the centromere proximal domain as well as the distal domain of major NORs can present different amounts of sequences, being the first case species-specific and the second case both individual (presenting heteromorphism) and species-specific.

Prophase silver stained chromosomes also revealed four distinct domains in these kind of Fagaceae NORs with two Ag positive bands flanking the secondary

constriction, as was already observed in *Brassica* spp. (Chen *et al*., 1995, Chen and Heneen, 1995). In addition simultaneous Ag staining and FISH revealed that these both domains are composed of condensed rDNA. We can therefore hypothesize that the proteins detected with silver belong to the transcriptional machinery of the rRNA genes but due to epigenetic marks or due to post-translational modifications of those proteins (Chen *et al*., 2005), or the combination of both, they remain silent until the cell requires the production of more proteins. FISH with 45S rDNA revealed in several nuclei two distinct chromatids emanated from condensed perinucleolar rDNA blocks indicating probably a lack of cohesin and condensin complex in this region which, in turn, can be related with transcription activity of these genes (Kobayashi and Ganley, 2005; Wang *et al*., 2006b).

Pseudosatellite is a special structure present in some nucleolar chromosomes of Fagaceae, and particularly evident in the terminal nucleolar chromosomes of *Castanea sativa.* From all the Fagaceae studied, *C. mollissima* is the only species that displays a true satellite without rDNA sequences. All the other Fagaceae studied, with the exception of *Fagus sylvatica*, have pseudosatellites in the major nucleolar chromosomes bearing rDNA sequences in a terminal position which present a special chromatin organization. In other species such as in *Brassica* spp. (Chen *et al*., 1995; Snowdon *et al*., 2002; Hasterock *et al*., 2006) and in *Beta* spp. (Desel *et al*., 2002; Dechyeva and Schmidt, 2006) where the rDNA *loci* appear in terminal position clear pseudossatellites can be detected although the traditional term satellite is used in those descriptions. Fagaceae pseudosatellite presents faint DAPI stain probably due to the richness in GC content of the 45S rDNA sequences, which in *Q. robur* is approximately 57%. Pseudossatellites can stay far away from the rest of the chromosome due to a huge secondary constriction resulting form high activity of rRNA genes what was probably responsible for some misinterpretations in previous cytogenetics studies where these domains were classified as supranumerary B chromosomes (Ohri and Ahuja, 1990, 1991).

A special NOR form was also observed in some plants with terminal NORs, the so called “peg” shape (Schroeder-Reiter *et al*., 2006). These structures analysed through scanning electron microscopy with 45S rDNA *in situ* hybridization revealed numerous small knobbly structures concealing parallel fibres, as opposed to the exposed parallel fibres of interstitial secondary constrictions deprived of knobs (Wanner and Formanek, 1995) indicating that a different chromatin organization is present in terminal ribosomal

chromatin. The knobbly structures of peg-like NOR seem to represent smaller chromomeres compared with those present in the chromatin of chromosome arms (Schroeder-Reiter *et al*., 2006). This distinct chromatin organization can also be responsible for the lighter DAPI staining observed in the pseudosatellite domain together with the high CG content as previously stated. Moreover, the large pseudosatellite of *C. sativa* proved to be composed of several different sequences and rearrangements. Sequencing of PCR products obtained with primers designed to amplify the complete IGS sequence revealed a small fragment flanked by both primers composed of a partial 26S rDNA mitochondrial fragment. Among flowering plants, nuclear genomes often contain significant DNA fractions of organellar DNA such as Nuclear MiTochondrial DNA (NUMT) which has been identified in almost all dicotyledons and monocotyledons (Richly and Leister, 2004 a, b) indicating that DNA sequence transfer from organelles to the nuclear genomes is a common process. Two main hypotheses try to explain how organellar DNA transfers to nuclear genomes: cDNA-mediated transfer and the bulk-DNA hypotheses (review Henze and Martin, 2001). The bulk-DNA hypothesis is currently the most accepted and suggests that gene transfer is a result of recombination promoted by a large amount of escaped organellar DNA, and is sustained by the presence of an entire mitochondrial DNA genome in the Arabidopsis nucleus (Lin *et al*., 1999; Stupar *et al*., 2001). According to this hypothesis, mitochondrial DNA can integrate into the nuclear genome by non homologous end-joining (NHEJ) repair (illegitimate repair) of double-stranded breaks (DSBs) (Yu and Gabriel, 1999; Ricchetti *et al*., 1999; Salomon and Puchta, 1998; Leister, 2005; Kuo *et al*., 2006). Repair of DSBs by NHEJ requires little or no sequence homology (0 to 4 bp, 'micro-identities') between the termini, enabling the noncomplementary ends of DSBs and organellar DNA to be 'pasted' to one another (Yu and Gabriel, 1999; Ricchetti *et al*., 1999). The *C. sativa* mitochondrial partial sequence is inserted in a reverse way sharing in the 5'end 7 bp with the 25S nuclear gene, and in the 3' end, 9 bp with the 18S nuclear sequence both in direct orientation (Fig 22a underlined). These small sequence homologies could be enough after a DSB to illegitimately repair two non homologous sequences, pointing to a direct DNA transfer that would have occurred in *C. sativa* genome in recent times since *C. sativa* NUMT share 98% of homology with mitochondrial sequences from other dicots. The NUMT sequence was physical mapped only in the pseudosatellites of *C. sativa*, although its location in the other NOR domains

can also be considered, as its detection in uncondensed chromatin could be under the resolution power of the FISH technique.

The PCR approach on Fagaceae genomic DNA with primers for the amplification of the complete IGS sequence also showed the existence of sequences shorter than the ones present in the basic tandem arrays revealed by southern blot hybridization in *Quercus* (Bellarosa *et al*., 1990). Sequencing analysis showed that these fragments contain a small part of the ETS with putative palindromic 18S rDNA sequences in both ends what can lead to distinct orientations within IGS sequences. A similar PCR approach used for studying genomic organization of subtelomeric repetitive sequences in rye revealed tandem arrays with units shorter than the ones detected by southern blots and with head-to-head and tail-to-tail disposition (Vershinin *et al*., 1995). Until recently rRNA genes were thought to be organized as a uniform head-to-tail tandem array however single-DNA-molecule analysis by molecular combing has revealed that human NORs comprise a mosaic of canonical and noncanonical rDNA repeats (Caburet *et al*., 2005). As many as one-third of human rDNA repeats are noncanonical forming palindrome structures separated by short IGS, while some canonical rDNA repeats are not oriented in telomere to centromere as previously reported. FISH with the Fagaceae palindromic sequences proved the co-localization with all domains of both major and minor NORs in *C. sativa* and to all 45S rDNA *loci* in *Fagus sylvatica*. This pattern is expected since the majority of the palindrome sequences constitute a fraction of the canonical repeat unit. Cycles of breakage-fusion-bridge first described by McClintock, 1941 can be responsible for the presence of palindromic sequences in Fagaceae's NORs. This mechanism has been attributed to the presence and amplification of a fraction of repeated subtelomeric units oriented head-to-head among head-to-tail units in rye (Vershinin *et al*., 1995). Dicentric chromosomes, resulting from asymmetric translocations, might initiate 'breakage-fusion-bridge'cycles, which involve repeated disruption and fusion of dicentric chromosomes during nuclear divisions (until the cycle is stopped by addition of telomeric sequences at the breakpoints) and cause duplications, deletions and/or inversions as secondary rearrangements in the chromosomes (review in Schubert, 2007). This mechanism can also be responsible for size heteromorphisms detected in homologous pseudosatellites. A similar mechanism to nucleoli fusion can be responsible for pseudosatellite fusion which is extensively observed in Fagaceae nuclei as well as in perinucleolar knobs in rye (Caperta *et al*., 2002), despite their distinct

chromosome location. Silver stained prometaphases of *C. sativa* (Fig. IV.9 e,f) revealed that pseudosatellites fuse either in opposite direction or can be perfectly superimposed what can be a consequence of the chromosome disposition in the previous interphase. Fusion in opposite direction can induce palindromic rearrangements in the rDNA sequences via somatic recombination what is corroborated by the increased number of palindromes sequences detected in human cells due to a defective helicase implicated in DNA replication and DNA repair pathways (Caburet *et al*., 2005).

DNA organization in several chromosomes domains has been correlated with DNA epigenetic marks and with differences in gene expression. Distinct epigenetic patterns were observed in different species of Fagaceae: in *F. sylvatica,* a species with small genome, there is a colocalization of the 45S FISH signal and the epigenetic marks associated with gene silencing (5-mC and H3K9me2) in all but one NOR *locus*, indicating that probably only one NOR is active. Moreover, this result is in accordance with the maximum number of nucleoli detected in meristematic root tip cells where only one small nucleolus was observed. Special domains in Fagaceae genomes such as pseudosatellites can display a high level of epigenetic marks associated with gene silencing being notorious the level observed in *C. sativa*.

*C. sativa* cells presenting heteromorphic pseudosatellites reveal the same methylation intensity in both pseudosatellites, suggesting that the amount of sequences present in this domain is not the only factor responsible for the high level of methylation detected. The same argument was also seen in the opposite situation where ecotypes with similar rRNA gene copy numbers show different NOR methylation levels in *A. thaliana* (Riddle and Richards, 2002). DNA sequences such as the ones identified in this work along with 45S rDNA sequences in the pseudosatellites of *Castanea sativa* may act as elicitors for methylation since palindromic segments of repetitive DNA sequences are associated with strong *de novo* hypermethylation in *Petunia hybrida* (Müller *et al*., 2002). Moreover, transcription of palindromic sequences can generate double stranded RNA activating the RNAi and siRNAs pathways which have been related with methylation in 68 *loci* of *A. thaliana* ecotype Ler but not in ecotype Col (Zhai *et al*., 2008). Small RNAs have also a role in the current model of NOR silencing involving the chromatin remodelling complex NoRC (reviewed in McStay and Grummt, 2008).

Taken together our results indicate that the chromosome structure pseudosatellite is not present in all genera of Fagaceae studied. Genus *Castanea* show distinct locations of rDNA *loci* since *C. mollissima* present a conventional satellite with the rDNA

sequences in interstitial position. Moreover, the pseudosatellite proved to be composed of several canonical and non-canonical rDNA sequences, and also of alien sequences such as the NUMT. This data suggest that in the evolution of the *Castanea* genus several genomic reorganizations have occurred leading to the prominent pseudosatellite of the European chestnut. We can speculate that a stronger presence in *C. sativa* NORs of "abnormal" sequences as the NUMT sequences and noncanonical rDNA repeats could be responsible for the strong heterochromatization and the enrichment of epigenetic marks associated with gene silencing in its pseudosatellites.

# V. Nuclear chromatin topology and distribution of epigenetic marks within the interphase nucleus

## V.1. Introduction:

The architecture of the nucleus depends on the nature and distribution of DNA sequences which are related to chromatin domains and gene expression, controlled and maintained by epigenetic processes, like cytosines methylation and histones tail modifications, and by protein interactions. Within the cell nucleus, many fundamental events involving physical alterations and movements of chromatin, such as enzyme complexes and promoters coming together turn chromatin accessible to polymerases (Heslop-Harrison, 2003).

Interphase chromatin organization is commonly associated with genome size being the proportion and distribution of different DNA sequences on chromosomes and chromatin domains directly related with genome size variation. In plants with a small genome size, and small chromosomes, like *Arabidopsis thaliana* (n=5; 157Mb/C (Bennett *et al*., 2003)), and *Medicago truncatula* (n=8; 466Mb/C (Arumuganathan and Earle, 1991)) most repetitive sequences, both in tandem (with the exception of rDNA and telomere repeats) as dispersed such as retrotransposons are localized preferentially in centromeric regions (Fransz *et al*., 2002; Kulikova *et al*., 2001). In large plant genomes, repetitive sequences, mostly LTR retrotransposons, are dispersed throughout the chromosomes, although some locations can be recognized as preferential (Pearce *et al*., 1996; Brandes *et al.,* 1997; Kumar *et al.,* 1997). Tandem repeats arranged in defined domains of the genome, can be microscopically recognizable as heterochromatic regions. The organization of constitutive heterochromatin in chromosomes and interphase nuclei vary according to genome size. A model for the interphase chromosomes organization of the small genome plant *Arabidopsis thaliana* (Fransz *et al*., 2002) proposes that chromosome territories are composed by heterochromatic chromocenters rich in repetitive sequences (including rDNA and centromere regions), from which gene-rich euchromatic loops emanate. On the other hand, in plants with large genomes characteristic euchromatic and heterochromatic regions can be less evident due to great dispersion of repetitive sequences. However in some species some heterochromatic blocks can be detected either subtelomerically (Vershinin *et al*., 1995) or interstitially (Siljak-Yakovlev *et al*., 2002) consisting of species-specific high repetitive satellite DNA. Interphase chromosome disposition in plants with large

genomes is characterized by a Rabl orientation where centromeres and telomeres are located at opposite nuclear territories (Morais-Cecílio *et al*., 1996), in opposition to small plants genome (Dong and Jiang, 1998). However, plants with intermediate genome size like maize (n=10; 2671 Mb/1C (Bennett and Smith, 1976)) displays neither entirely Rabl nor non-Rabl organization. Also, in some species Rabl organization can vary among tissues or developmental stages of an organism (review Cowan *et al*., 2001).

Associated with heterochromatin or euchromatin are several epigenetic marks such as the posttranslational modification of histones, and the methylation of the cytosines. The nuclear distribution of heterochromatin associated marks like the conserved methylated histone isoforme H3K9me2 in plants and the hypermethylation of cytosines can differ according to the organization pattern of repetitive sequences, and therefore to genome size. In Arabidopsis these marks are restricted to the chromocenters (Fransz *et al*., 2002; Houben *et al*., 2003) while in large plants genome there is a more uniform distribution throughout the nucleus (Castilho *et al*., 1999; Houben *et al*., 2003).

The Fagaceae genome organization was studied through the isolation, characterization and physical mapping of various repetitive sequences as well as through the detection of several epigenetic marks patterns.

## V.2. Materials and Methods

### V.2.1. Plant material:

Root-tips were collected from germinated acorns or from seedlings in pots from the following species: *Fagus sylvatica* L., *Quercus suber* L., *Quercus ilex* L. subsp. *rotundifolia* (Lam) O. Schwarz, *Quercus acutissima* Carruth., *Quercus serrata* Murray, *Castanea crenata* Sieb. & Zucc d), and *Castanea mollissima* Bl.

### V.2.2. AFLP FISH probes:

Genomic DNA from *Quercus ilex* subsp. *rotundifolia* was first digested with *Eco*RI and *Mse*I restriction enzymes, and further ligated to adaptors and amplified using two different combinations of *Eco*RI and *Mse*I primers with three selective bases as described in section **II.2.2**, using the AFLP core reagent kit (Invitrogen). The resulting AFLP fragments were detected in a denaturing polyacrylamide gel stained with silver nitrate. Several fragments with lengths ranging from 100 to 500 pb were excised from the gel. FISH probes were produced by PCR amplification of each fragment using the primer combination previously utilised to generate them, along with digoxigenin labelled nucleotides.

### V.2.3. Sequencing and homology search

Most fragments were directly sequenced from PCR products, while others were cloned using the TOPO TA cloning kit (invitrogen), prior to sequencing. Sequences homologies search was performed using either BLASTn and BLASTx tools (Altschul *et al*., 1997).

### V.2.4. Nuclei preparations:

For fluorescence *in situ* hybridization and immunodetection of methylated cytosines, fresh ethanol: glacial acetic acid (3:1 v/v) was used as fixative to preserve the plant material, and preparations were done according to the drop technique described in section **II.3.2**. *In situ* immunolocalization of modified histones was performed in roots fixed in fresh formaldehyde solution 4% (w/v), and slides were obtained according to the method described in section **II.3.3**.

### V.2.5. DNA:DNA FISH

Five AFLP-probes labelled with digoxigenin and detected with anti-digoxigenin antibody conjugated to FITC, giving a green signal with the appropriate microscope filters set were used as described in section **II.5**.

### V.2.6. Fluorescent *in situ* immunodetection

The following three primary antibodies were used separately: anti-5mc raised in mouse (Abcam AB 10805); anti-H3K4me3 (Abcam AB 8580) and anti-H3K9me2 raised in rabbit (Abcam AB 7312). All secondary antibodies were conjugated with Cy3 fluorochromes, giving a red signal with the appropriate microscope filters set. The immunodetection of the modified histones and methylated cytosines followed the procedures described in sections **II.6.2** and **II.6.3**, respectively.

### V.2.7. Cell analysis and image acquisition:

Fluorescent slides were analysed with an epifluorescence microscope (Axioskop 2, Zeiss). All images were collected using an AxioCam digital camera (Zeiss) controlled by AxioVision 3.0 and assembled using Adobe Photoshop 6.0.

## V.3. Results

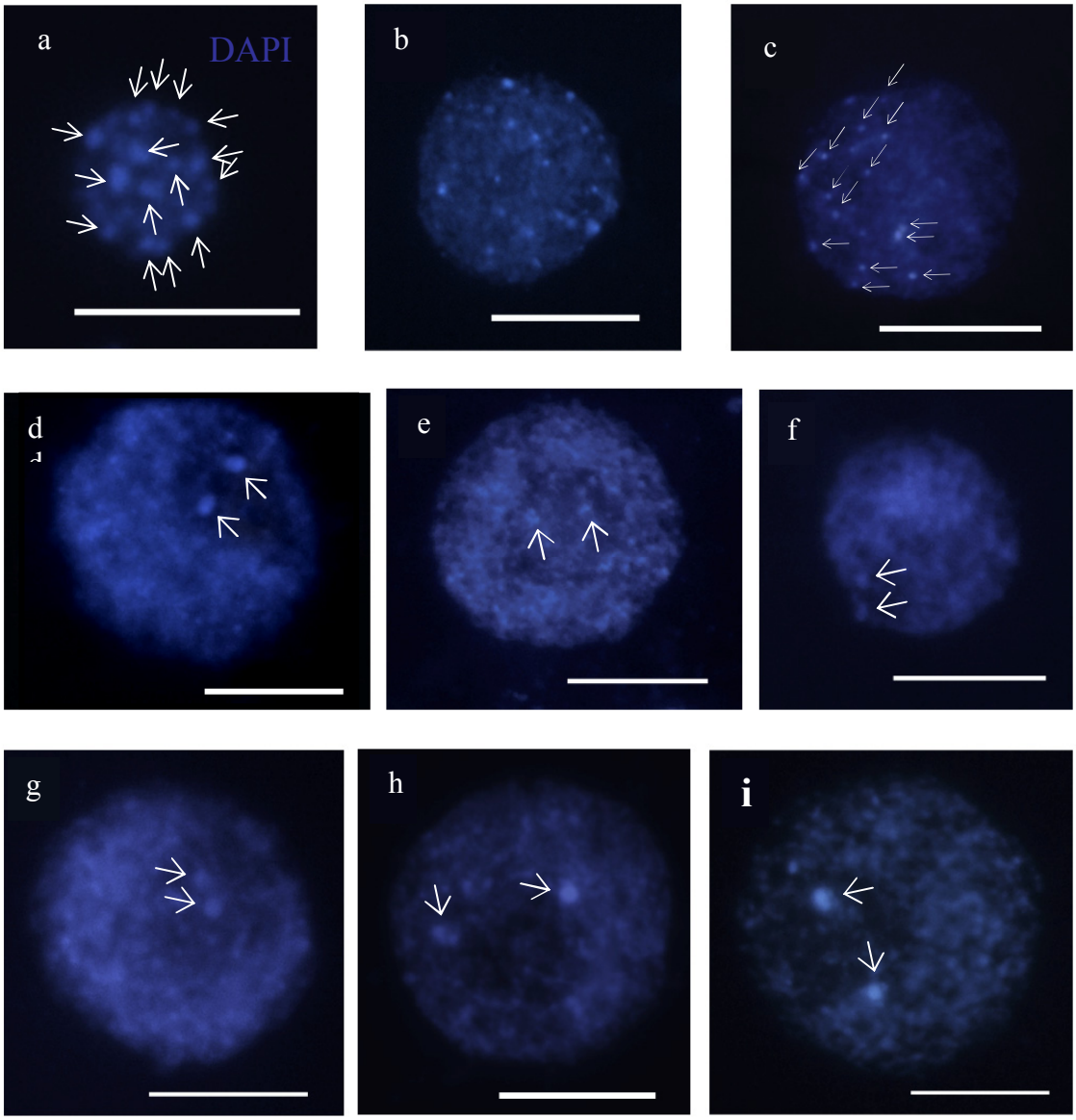
### V.3.1. Chromatin organization in Fagaceae interphase nuclei

Heterochromatic and euchromatic interphase regions were evaluated trough DAPI staining in different members of the Fagaceae family which include species of the three genera in study: *Fagus*, *Castanea* and *Quercus* (Fig. V.1).

The species studied revealed different patterns of interphase chromatin organization varying from species with well defined chromocenters (Fig. V.1 a,b,c) to species with evenly distributed heterochromatin regions (Fig.V.1 d-i). Interphase nuclei of *Fagus sylvatica* present clear chromocenters. The average number per nucleus observed is 29 ± 7 (*n*=25) when the maximum expected number is 32 (corresponding to the 24 centromeric domains (Fig. IV.2 a) + 8 45S rDNA *loci*). Differences in intensity can be observed being the rDNA blocks (inferred from Fig. IV.10) the more intense DAPI stained, and the smaller pericentromeric ones the faintest (Fig.V.1 a, white arrows). Large DAPI blocks, probably resulting from associated smaller chromocenters can be detected in nuclei with chromocentre numbers less than the maximum expected (Fig. V.1 a). Nuclei of *Quercus serrata* (Fig. V.1 b,c) showing very well defined small chromocenters with similar sizes differ from those of others *Quercus* species studied (Fig. V.1 d-f). Interphase nuclei of *Quercus serrata* show an average of 23 ± 4 chromocenters (*n*=50) when the maximum expected number is 26 (corresponding to the 24 centromeric regions (Fig. I.6 c inset) + two 45S rDNA perinucleolar knobs). One third of the nuclei observed present the chromocenters located at one nucleus pole (Fig. V.1 c) resembling a Rabl configuration, which was not found in *Fagus sylvatica*. In the other *Quercus* and *Castanea* species studied (Fig. V.1 d-i), interphase nuclei show a more uniform chromatin structure and no clear chromocenters can be seen excepted for one or two heterochromatic knob(s) located at the periphery of the nucleolus (white arrows) corresponding to the major rDNA repeats (Fig. V.1 d-i) as inferred by the data present in chapter IV.

**Figure V.1** – Meristematic root-tip interphase nuclei of *Fagus sylvatica* **a)**, *Quercus serrata* **b,c)**, *Quercus ilex* subsp. *rotundifolia* **d)**, *Quercus suber* **e)**, *Quercus acutissima* **f)**, *Castanea sativa* **g)**, *Castanea crenata* **h)**, and *Castanea mollissima* **i)** stained with DAPI. Chromocenters **(a,c)** and major perinucleolar rDNA heterochromatic knobs are indicated by white arrows **(d-i)**. *Q. serrata* nucleus **(c)** presenting chromoceters located at one pole. Bars=10 μm.

**Figure V.1**



### V.3.2. Distribution of epigenetic marks within the somatic interphase nucleus

Patterns of epigenetic marks associated with active (H3K4me3) or inactive (5-mC and H3K9me2) chromatin were studied in meristematic and differentiated interphase nuclei of species presenting different interphase chromatin organization: *Fagus sylvatica* having large chromocenters, and *Q. suber, Q. ilex* subsp. *rotundifolia* and *C. sativa* showing a more uniform interphase chromatin organization. In all species studied, the epigenetic patterns were similar in nuclei from meristematic and differentiated cells.

Immunodetection of 5-mC was performed in ethanol-acetic acid fixed nuclei, while detection of modified histones was carried out after formaldehyde fixation. The type of fixative accounts for the differences detected in DAPI stained nuclei, giving the ethanol-acetic fixative a better chromatin definition

Detection of 5-mC and H3k9me2 revealed strong signals on the large chromocenters of *Fagus sylvatica*, and is also weakly present throughout the nuclei. H3K4me3 showed disperse signals without any particular region of concentration, being however absent from the chromocenters (Fig. V.2 g-i, white arrows).

In *Castanea sativa*, one of the non-chromocenter type species, heterochromatic epigenetic marks show a spotty pattern having a global distribution of the 5-mC labelling, while the H3K9me2 pattern is present in larger signals, although there is a strong concentration of 5-mC and an enrichment of H3K9me2 in the perinucleolar knobs (Fig. V.3 d,e) corresponding to the pseudosatellites. The H3K4me3 pattern is uniform throughout the entire nuclei (Fig. V.4 f).

The immunodetection of epigenetic marks in *Q. suber* and *Quercus ilex* subsp. *rotundifolia* revealed similar patterns, therefore only *Q. suber* nuclei are shown (Fig. V.4). In *Q. suber* the immunodetection of epigenetic marks revealed the same type of pattern described for the *Castanea sativa* except for higher concentration of 5-mC and the enrichment of H3K9me2 in the perinucleolar knob. The 5-mC distribution pattern was also studied in a triploid genome of *Q. suber* in order to evaluate potential alterations of the diploid pattern. Meristematic interphase nuclei show the same pattern observed in diploid genomes with a conspicuous label all over the nucleus (Fig. V.5 a), demonstrating that there are no DNA methylation marked differences between the

diploid and the triploid condition. Metaphase chromosomes (Fig. V.5 b) present an odd distribution with regions intensely labelled and others with no label. All complements sets share the same distribution and intensity of signal.

**Figure V.2**- Epigenetic marks immunolocalization in *Fagus sylvatica* differentiated (**a**) and meristematic (**d**,**g**) root-tip cell nuclei. Chromatin counterstained with DAPI (**a,d,g**); 5-mC immunolocalization (**b**); H3K9me2 immunolocalization (**e**); H3K4me3 immunolocalization (**h**); Superimposition of images (**c**, **f**, **i**). The chromocenters that lack H3K4me3 immunollabeling are indicated by white arrows. Bars= 5 μm.

**Figure V.3**- Epigenetic marks immunolocalization in *Castanea sativa* meristematic (**a**,**g)** and differentiated (**d**) root-tip cell nuclei. Chromatin counterstained with DAPI (**a,d.g**); 5-mC immunolocalization (**b**); H3K9me2 immunolocalization (**e**); H3K4me3 immunolocalization (**h**); Superimposition of images (**c**, **f**, **i**). Bars= 10 μm.

**Figure V.2**

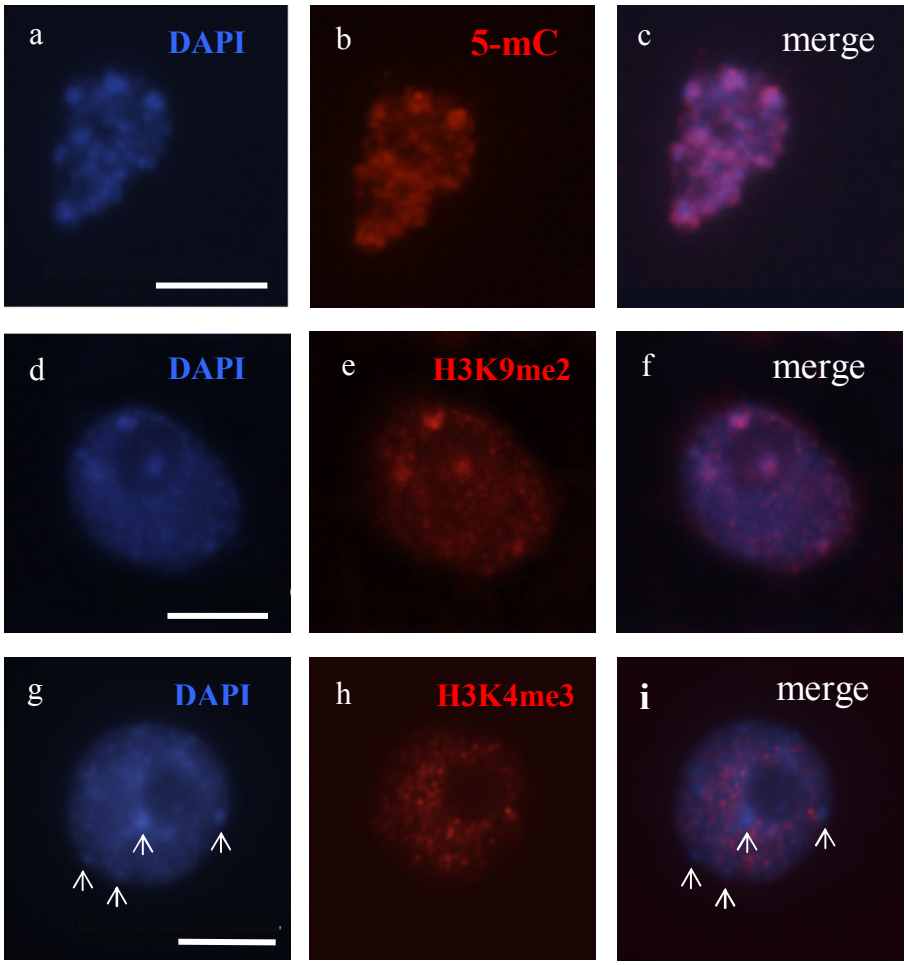



**Figure V.3**

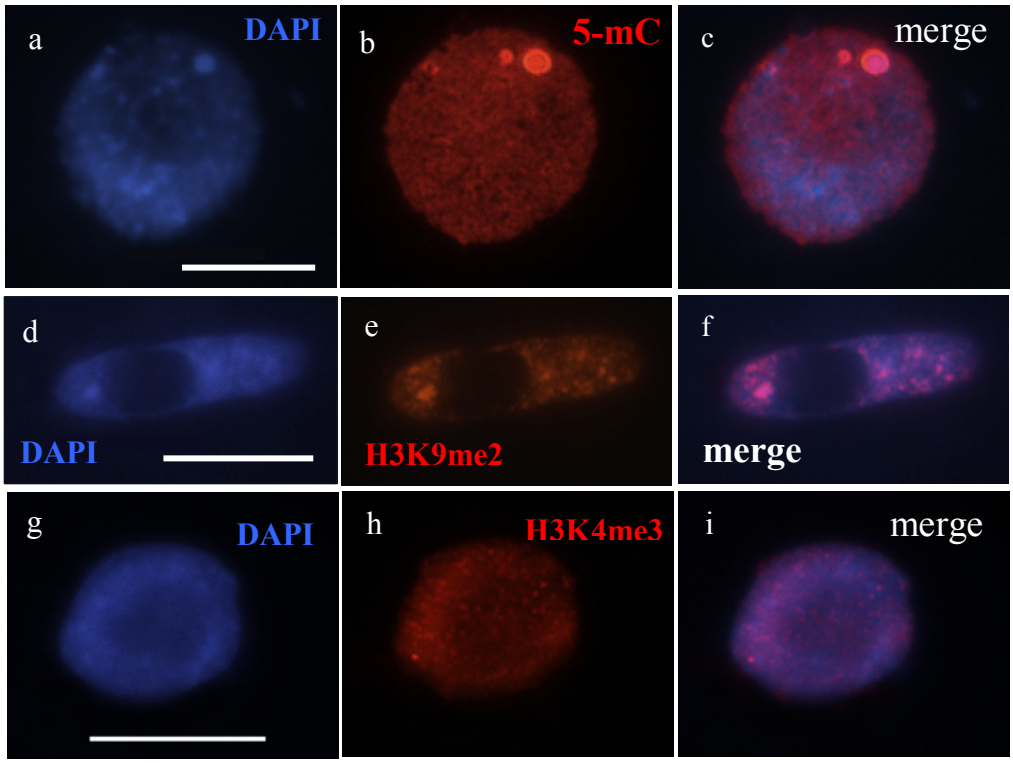


**Figure V.4**- Epigenetic marks immunolocalization in *Quercus suber* meristematic root-tip cell nuclei. Chromatin counterstained with DAPI (**a**, **d**, **g**); 5-mC immunolocalization (**b**); H3K9me2 immunolocalization (**e**); H3K4me3 immunolocalization (**h**); Superimposition of images (**c**, **f**, **i**). Bars= 10 μm.

**Figure V.5** –Methylated cytosines immunolocalization (red) in *Quercus suber* triploid root-tip meristematic nucleus **a)**, and prometaphase chromosomes **b)**. DNA counterstained with DAPI. Bars=10 μm.

**Figure V.4**

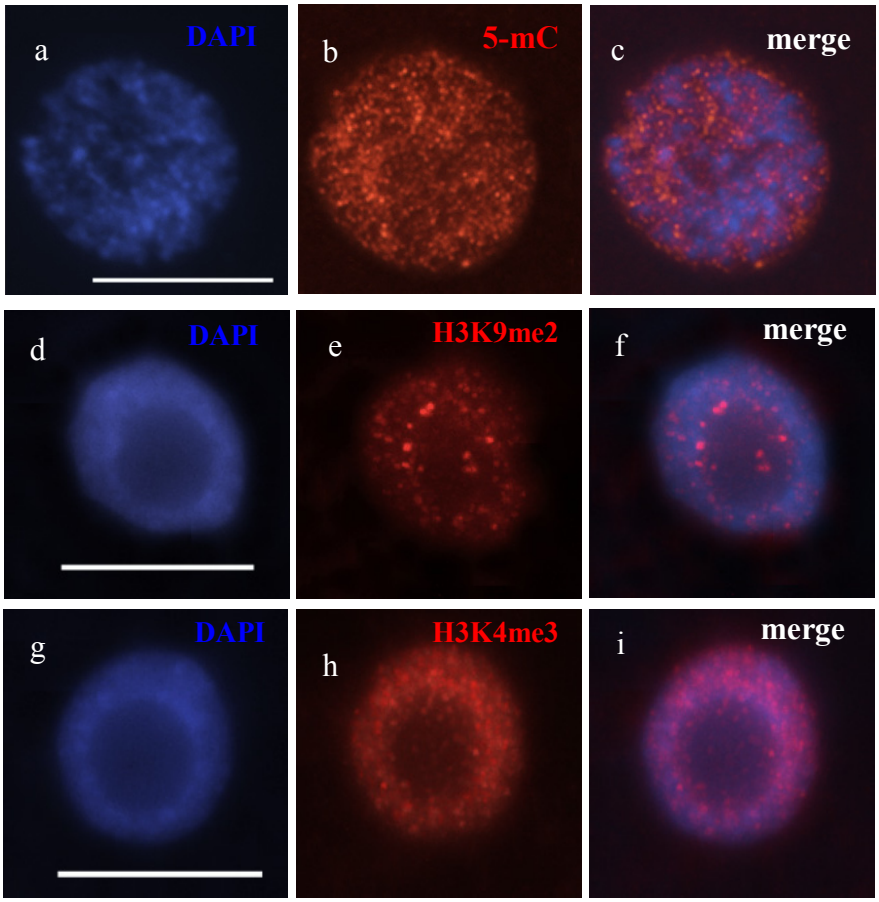
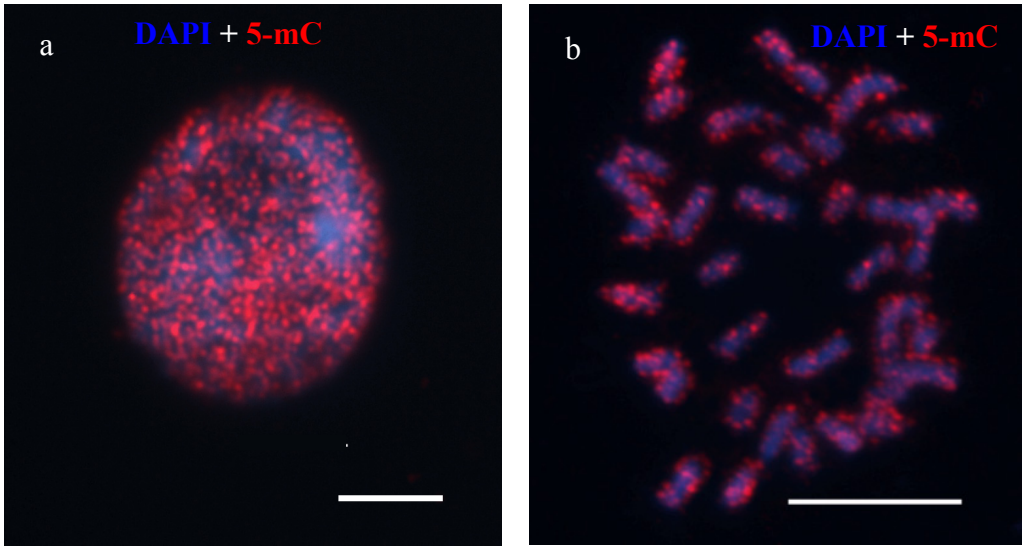



**Figure V.5**



### V.3.3. Isolation of repetitive sequences and their distribution in the Fagaceae nuclei

The isolation of repetitive sequences specific of members of the Fagaceae family was performed using the AFLP technique on *Q. ilex* subsp. *rotundifolia* genomic DNA. AFLP DNA fragments were used to produce FISH DNA probes in order to study their distribution patterns in several members of the Fagaceae.

AFLP fragments correspond to unique positions in the genome and hence can be exploited as landmarks not only in genetic but also in physical mapping. In order to isolate repetitive sequences, AFLP derived fragments were produced with two different combinations of *Eco*RI (E) and *Mse*I (M) primers, ending in different nucleotides: combination1=E-AA and M-CAA and combination2=E-AC and M-CAG. The amplified products were separated through denaturing polyacrylamide gel electrophoresis and further stained with silver nitrate (Fig. V.6). Five fragments with lengths between 100 to 500 bp and exhibiting strong labelling were chosen to be characterized. Sequencing of these fragments revealed that their sizes ranged from 132 to 517 bp (Table V.1). Moreover, all sequences have an high AT content, being the highest amount observed in Rot 18 sequence with 61.5% , and the lowest in Rot 10 with 51.5% (Table V.1).

Nucleotide homology search was performed with the five sequences using the BLASTn algorithm with default parameters. Rot 2 presents 78 to 80% of homology in 85 to 60% of its length with three genomic sequences: two of *M. truncatula* located on chromosome 3 (CT967305.12, CT967316.8), and one from *Vitis vinifera* (AM467565.1), with very low expected values (e-values) ($1^{e-06}$ from nucleotide 4 to 116, $5^{e-05}$ from 34 to 116 and $4^{e-07}$ from 39 to 118, respectively).

Also Rot 20 and Rot 10 show homology with *V. vinifera* (AM446457.2) and *Phaseolus vulgaris* (DQ323045.1), respectively with low e-values (from $2^{e-05}$ to $3^{e-05}$), although in only 15% and 6% of its length, from nucleotide 283 to 344 and from 220 to 252, respectively. The other two Rot sequences, Rot 8 and Rot 18, failed to show homology with any sequence of the databases.

Also a search for sequences homology at the predicted amino acid level against protein databases was performed using the BLASTx algorithm. Rot 10 show high homology with several sequences related to retrotransposons. All the retrieved

sequences in the search are retrotransposon related proteins or hypothetical proteins, and the expectation values ranged from $4^{e-26}$ to $2^{e-21}$ being the first records the following: 38% identity, 61% positives and an expectation value of $4^{e-26}$ with a pol-polyprotein from a gypsy-like retrotransposon from *Silene latifolia* (AAY99339.1 GI:68685649); 39% identity, 59% positives and an expectation value of $2^{e-25}$ with a putative pol-polyprotein of *A. thaliana* (AAD20433.1 GI:4417309); and 38% identity, 56% positives and an expectation value of $5^{e-24}$ with a retrotransposon protein putative Ty3-gypsy subclass from *Oryza sativa* japonica cultivar- group (AAP52584.2 GI:110288766). The homology search of the putative protein sequence of Rot 10 in the conserved domains databases with default parameters (Marchler-Bauer *et al*., 2007) produced matches with RNase H superfamily with an e-value of $1^{e-11}$. This protein was associated not with a particular RNase H domain but with a superfamily, which represent evolutionarily related domains, conserved domains shared for example by *Sorghum bicolor* (gi:75215985) and *Zea mays* (gi:75315872).

The Rot 20 sequence also shows significant hints using BLASTx algorithm, although at the predicted amino acid level two sequences were generated corresponding to two different frames (-1 and -2). The homology with the longest amino acid sequence (90 out of 130) show 48% identity, 72% positives and an expectation value of $7^{e-25}$ with a reverse transcriptase family protein from *Asparagus officinalis* (ABD63139.1 GI:89179403 ) and within 96 amino acids 47% identity, 71% positives and e-value of $3^{e-24}$ with RNase H family protein from the same organism (ABD63162.1 GI:89179427). However, search for conserved domains failed to give any result with default parameters.

The five sequences studied were PCR amplified and labelled in order to produce FISH probes to physical mapping. All the sequences hybridize with one or several species of the Fagaceae family, but not with all members of the entire family (Table V.2, Fig. V.7). Rot 10 is present in all *Quercus* spp. analysed (European and Asiatic) and in *Castanea sativa*, (Fig. V.7 a-e) but not in *Fagus sylvatica*, while Rot 20 is present only in the genus *Quercus*; both *Castanea sativa* and *Fagus sylvatica* did not show hybridization signals (Table V.2). These two sequences are dispersed in the genomes, in agreement with the pattern displayed by retrotransposons present in medium/large genomes. The other three sequences (Rot 2, 8 and 18) are *Quercus ilex* subsp. *rotundifolia*-specific. As these sequences are not present in *Quercus suber* (a closely related sympatric species) no other species were tested. Rot 2 and Rot 8

sequences, show an even distribution throughout the nuclei with no preferential cluster sites. From all the sequences studied, only Rot 18, the sequence with the highest AT content has a different pattern being detected in a cluster-like organization at stronger DAPI regions, although not forming chromocenters (Fig. V.7 j).

**Table V.1** – Nucleotide sequence and AT/GC content of AFLP derived fragments used as FISH probes

| | % | | | |
|---|---|---|---|---|
| **FISH probe** | **AT** | **GC** | **Sequence** | **No. nucleotides** |
| **Rot2** | **57.5** | **42.5** | AATTCCCCCATATNACCGGTCTTCCTGGAGA TCAATTAGCTGATGGGTTTTGCGGCGCAAAA CTTGAATTTGGGATAGTACAGCTTTTCATTAT TGAGTTTCATGAAGACTGTTGTTACTCAGGA CTCATCA | **132** |
| **Rot18** | **61.5** | **38.5** | GCTTGTGTCCTGTTTTTATGTGCATAACACAC AAGCTAAACCAAGTGAAAAGCTGCCACATAT TTATAAAGAGACTTCCATGAAAATTCCATAA ACCATAATTGTGTTTCATTTTGGTTGGTGACA GTGGGTCATGTGTACTCAGGACATCACCACC ACG | **160** |
| **Rot8** | **58.9** | **41.1** | CAGGTAGTTTCTTATTCCTCGTCAGTTCCTTT GGCTCCCTTGATCATGTGATATGGTACCWKA ACATCTTGACTAGATTTGTCGAAGTTTTTCCT GCCCAGGATCAAAACCCGAGCATACATTTAT ATGARAGCTMATGAWGACATACAAACGTAT ATGCAGGCTTGKKGGATAAAAAGGCATTCCC ATCAACTGCAAAGACTGTTGTTACTCAGGAC TCATCAA | **225** |
| **Rot20** | **55.1** | **44.9** | AGGTGGTGCAATATTAGGCAGCCGGGAACTC ATGCTTTTGCGGACATTCCTTGTTTAGGTTGC AAAAGTCCACACAGCAACGGATTTGGCGGTT CTTTGTTCTTCATAGGTACTATGTTAGAAAGC CATTTGGGATGTTGAATGGACTTGATAAAAC CAGCTGCTAGCAACATCTTGACTTCTTGAGTT ATCTGAGTTTCTACATCAATGTGAAAGACCT TGGCTGGTTGGATTACTGGCTTGACTCCTGG GTCCACATTGAGAGCATGGACCACCAGCCTT GGATCCAGACCAAGCATCTCATCATAGGTCC ATGGGAACACATCCATGTATTCCTTGAGCAA GGCCACTAATTGCTCCTTTTCTTATGCTGTTA CTCAGGAACTTCATCA | **392** |
| **Rot10** | **51.5** | **48.5** | TTCAACGTCCAATACCGCCCGCATACTACCA TGAAGGGACAAGCAGTCGCTGACTTTATTGC GGAATTCACCAATATGGAAGGCCAGGGGGC AGAAGAGCATCCTCAATGGAGTATCCACGCG GATGGATTGTCCAACAAGCAAGCTTGCGGAA AAGGTATAGTACTCCACTCACTAGAAGGGGA TGAGATTGAGTGCATGGTTCATCTTGACTTCC CTACGACCAACAATGAAGCGGAGTACGTGG CTCTAGTGGTAGGACTGGATCTCGCCAAAGC AGTAGGGGCCATATGTGTGGTTGTGTATTAC AACTTTTAGGTGGTCACAAGTCAGGTGAACG GTGACTACGAGTGCAAAGGTGAAAGGATGA AGAAATACTTGAAGCAAGTAAGGAAGCGGG TGGGTGACCTTCAGGTCAGGTTTGTTCAAAT CCCAAGGGAAGAGAATAAGCAAGCTGACTG CCTTGCCAGAGCTGCATCAGCCGAACATATG CTCATCCCCAATAAGGTACTTTCTT | **517** |

**Figure V.6** - Polyacrylamide gel silver stained revealing the AFLP fragments generated from *Quercus ilex* subsp. *rotundifolia* genome, with two combinations of selective *Eco*RI (E)/*Mse*I (M) primers, ending in different nucleotides with two replicates: P1 - Primer set1 combination=E-AA and M-CAA, and P2 - Primer set2 combination=E-AC and M-CAG. Fragments excised from the gel and further characterized are indicated by arrows.

**Figure V.6**

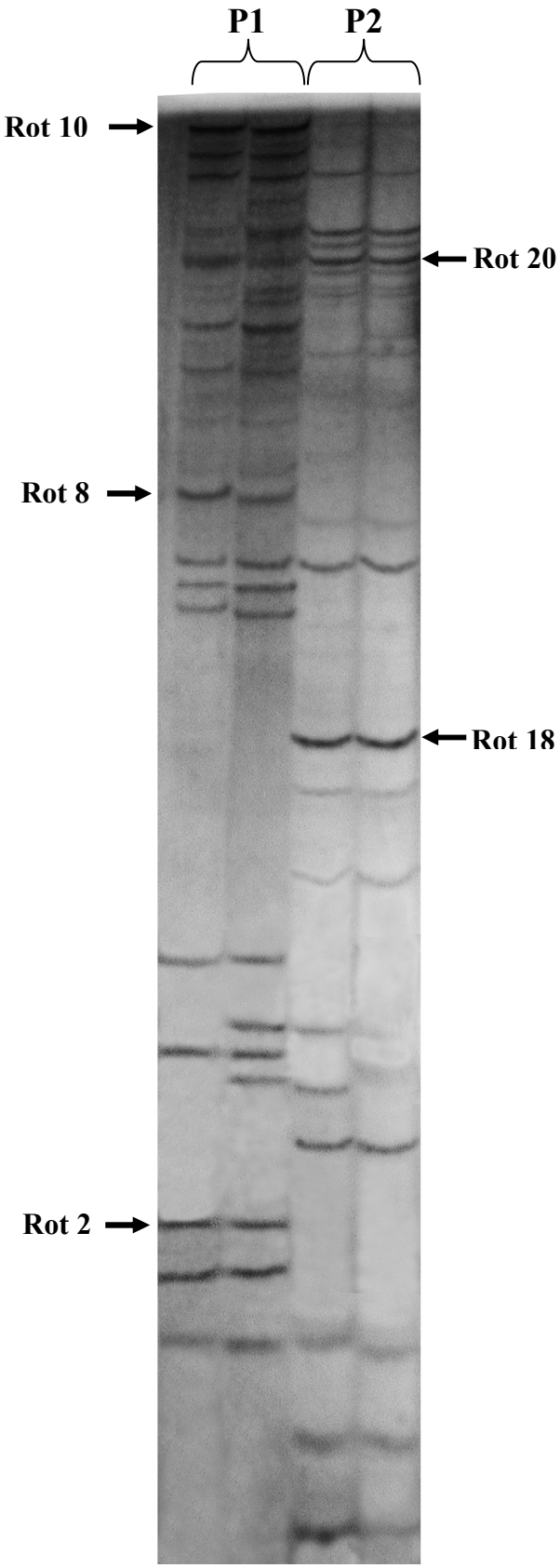


**Figure V.7-** FISH with AFLP derived sequences in meristematic interphase nuclei of *Quercus ilex* subsp. *roundifolia* **a),f),h-j)**; *Quercus suber* **b),g)**; *Quercus acutissima* **c)**, *Quercus serrata* **d)** and *Castanea sativa* **e)**. Probes were labelled with digoxigenin and visualized with α–digoxigenin conjugated with FITC. Nuclei are counterstained with DAPI.

**Table V.2** –FISH hybridization of different AFLP derived sequences in distinct Fagaceae species.

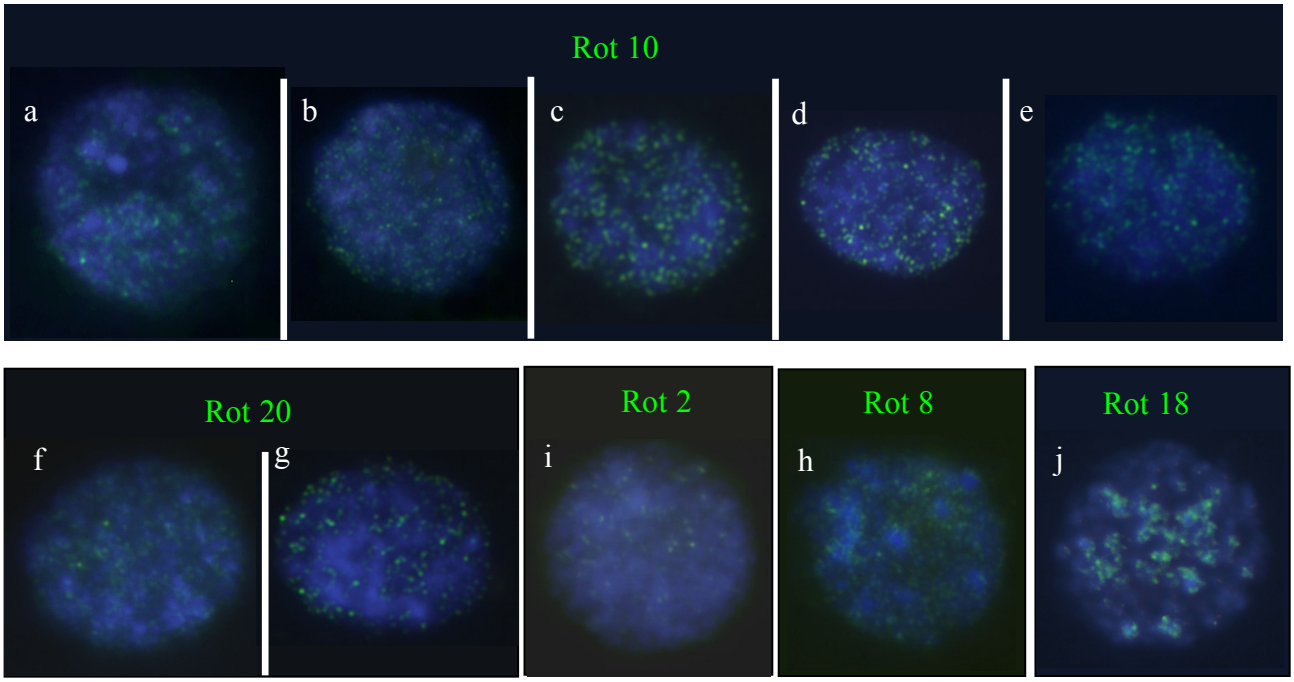
| | *Q. ilex* | *Q. suber* | *Q. acuttissima* | *Q. serrata* | *C. sativa* | *F. sylvatica* |
|---|---|---|---|---|---|---|
| Rot 10 | + | + | + | + | + | - |
| Rot 20 | + | + | | | - | - |
| Rot 2 | + | - | | | | |
| Rot 8 | + | - | | | | |
| Rot 18 | + | - | | | | |

+ positive hybridization

- negative hybridization

The lack of marks represents the absence of hybridization experiments.

**Figure V.7.**



## V.4. Discussion

In plants interphase nuclei were classified into three types according to the heterochromatin distribution: nuclei with a diffuse heterochromatin pattern associated with large genomes; nuclei with a heterochromatin gradient displaying large masses of interphase heterochromatin, also associated with large genomes; and nuclei with marked chromocenters as in *Arabidopsis*, typical of small genomes (Fukui and Nakayama, 1996 reviewed in Reuter *et al*., 2005). For the gradient organization, blocks of heterochromatin are frequently located in one half of the nucleus and are associated with species with a typical Rabl orientation, like *Secale cereale* and *Vicia faba* (Reuter *et al*., 2005). According to this classification, besides *Quercus serrata* and *Fagus sylvatica* which present chromocenter type nuclei, all the Fagaceae studied have an interphase heterochromatin distribution of the diffuse type, being the perinucleolar rDNA knob the most recognizable heterochromatic fraction markedly enriched with silent chromatin associated epigenetic marks as in *Castanea sativa*. Although the classification of diffuse heterochromatin seems contradictory with the traditional concept i.e. a portion of chromatin that stays constitutively or facultatively condensed all over the cell cycle, we assume the diffuse organization on nuclei with the heterochromatin fraction evenly distributed, and with no significant cytologically visible concentrations. Immunolabeling of modified histones in species of *Quercus* (except *Q. serrata*) and *Castanea* points to nuclei with heterochromatin of the diffuse type. This organization is in agreement with heterochromatin fraction intermingled with euchromatin since H3K4me3, H3K9me2 and 5-mC appear evenly distributed all over the genome.

The 5-mC epigenetic pattern of diploid and the triploid *Q. suber* genomes was compared. Although *in situ* immunodetection is a broad-scale technique and fine-scale epigenomic studies will be essential to deep our knowledge on triploid trees epigenetics, no visible effects seem to be induced by the addition of a whole genome. The balance hypothesis proposes that regulatory genes are retained in duplicate because a change in their relative number modifies the expression of target genes (Birchler *et al*., 2005). Therefore one would think that in the triploid individual of *Quercus*, the extra complete set of regulatory genes and of genes whose products participate in protein–protein

interactions would be silenced. In Arabidopsis it was shown that the transgene expression in plants with different ploidies is altered indicating that the ploidy level can affect epigenetic silencing (Scheid *et al*., 1996). However the level of changes in gene expression detected in autopolyploids is much lower when compared with changes verified in allopolyploids (Wang *et al*., 2006b), where gene silencing is known as a common response (Kashkush *et al*., 2002; He *et al*., 2003; Adams *et al*., 2003; Adams *et al*., 2004) mediated by epigenetic mechanisms and therefore related with changes in CpG methylation (Lukens *et al*., 2006). Therefore the *in situ* immunodetection of 5-mC may not detect possible epigenetic pattern differences when triploid and diploid *Q. suber* genomes are compared.

*Fagus sylvatica* although having evident chromocenters does not present the immunolabelling pattern characteristic of the chromocenter type nuclei since the heterochromatin epigenetic marks although stronger at the heterochromatic blocks also appear disperse all over the nuclei as in *Scilla mischtschenkoana,* a species with gradient type nuclei (Fisher *et al*., 2006). This pattern indicates that highly represented repetitive sequences like transposable elements are probably distributed along chromosomes in *Fagus sylvatica* instead of being exclusively located at the centromeric region. This chromatin distribution could explain the immunolabeling patterns detected since for instance, DNA methylation and H3K9me2 all over the nuclei would be necessary to silence retroelements (Richards and Elgin, 2002) and prevent their movement and expansion next to genes. Interphase chromocenter disposition located at one pole in *Quercus serrata*, points to an interphase chromosome Rabl orientation which, in turn, has been related with large genomes having nuclei of the gradient type. These patterns points to a different interphase chromatin organization different from the typical chromocenter organization type and from the "diffuse" type like the other *Quercus* and *Castanea* spp. analysed. Therefore, *Fagus sylvatica* and *Quercus serrata* cannot be included in the classification proposed by Fukui and Nakayama (1996) and reinforced by Reuter *et al*. (2005) indicating that other heterochromatin interphase organizations are possible in plant genomes.

Our finding also clearly contradicts previous works that tried to establish a solid correspondence between genome size and heterochromatin interphase organization (Houben *et al*., 2003). Indeed our results clearly show that different *Quercus* spp. with similar genome sizes and the same chromosome number present distinct interphase patterns of chromatin organization, as already suggested by Zoldoš *et al*. (1999) for

*Q. cerris* (931 Mb/1C (Zoldoš *et al*., 1998)) and *Q. suber* (931 Mb/1C (Zoldoš *et al*., 1998). Moreover, the large genome species *Ornithogalum longebracteanum* (7448 Mb/1C (Bennett and Smith, 1976)) have a typical Arabidopsis-like heterochromatin interphase organization (Reuter *et al*., 2005; Fisher *et al*., 2006) which corroborate our findings on the absence of a direct correlation between genome size and chromatin organization. This finding also seems to indicate that the nature of repetitive sequences and their preferential distribution are determinants to nuclear heterochromatin interphase organization of plant species.

*Quercus ilex* subsp. *rotundifolilia* with a DNA content of 970 Mb/1C (Ribeiro and Loureiro, unsubmited data) is probably the Fagaceae genome with more repetitive sequences from the species studied: *F. sylvatica* (544 Mb/1C (Gallois *et al*., 1999)); *Q. suber* (931 Mb/1C (Zoldoš *et al*., 1998)); *C. sativa* (956 Mb/1C (Barow and Meister, 2002)). Repetitive sequences derived from AFLP fragments have been used as FISH probes (Reamon-Buttner *et al*., 1999) rendering AFLP an efficient technique for producing cytogenetic molecular markers, besides its importance for DNA fingerprinting (Vos *et al*., 1995). We characterized in *Q. ilex* subsp. *rotundifolilia* five AFLP fragments with strong intensity, with sizes varying from 100 to 500 bp suitable for good penetration in FISH procedure (Schwarzacher and Heslop-Harrison, 2000) and their interphase distribution pattern was further evaluated.

Those five sequences isolated from *Quercus ilex* subsp. *rotundifolia* genome present moreover different contribution to the family Fagaceae, since Rot 10 is present in two close genera *Quercus* and *Castanea*; Rot 20 is genus-specific; while Rot 2, 8 and 18 seem to be species-specific. All the sequences analysed but one have dispersed distribution patterns characteristic of transposable elements in medium and large genomes (SanMiguel *et al*., 1996), although bioinformatic analysis only confirms a marked retroelement homology for Rot 10 sequence and a smaller homology for Rot 20. Combining bioinformatics analysis on predicted protein and conserved domains indicates that Rot 10 belongs to an RNase H gene, which codes for an enzyme that specifically degrades the RNA present in a RNA/DNA heteroduplex of Ty3-gypsy LTR-retrotransposon. Retroelements contribute greatly to plant genomes constitution being the LTR retrotransposons with two major subclasses (g*ypsy* and *copia*) the most abundant (Rabinowicz and Bennetzen, 2006). Despite their common components, retroelements are very heterogeneous and, because they evolve faster than other gene sequences, are often found to be species specific (Heslop-Harrison, 2000), although the

same types of elements can be found in different plant species. The presence of Rot 10 retroelement-like sequence in only two related genera (*Quercus* and *Castanea*), but not in the more distant *Fagus* indicates that the diverging time between the two close genera was not enough for this retroelement to evolve towards a distinctive type. Although the Rot 20 sequence shows some homology to retrotransposons at the predicted protein level, it did not reflect a considerable retrotransposon conserved domain. The other Rot sequences failed to show any homology related with retrotransposons with the entire database although they also present a disperse FISH pattern. Transposable elements appear to change their own structure within the same genome, deletions and other internal rearrangements are common, perhaps as an outcome of failed transposition events, resulting in unrecognizable sequences (Ma *et al*., 2004). Rot sequences that failed to display homologies with the databases could be putative derivatives of retrotransposons that have pass for this type of process becoming unrecognized or could simply be completely different types of repetitive sequences.

Although Rot sequences present sizes below FISH resolution, they are clearly seen in interphase nuclei probably revealing the way they are organized in small tandems, or in the case of Rot 10 and Rot 20 how these retroelements or now possible derivatives (Rot 20) have integrated in the plant genomes, probably resulting from several insertions events into each other, as has been described for many retroelements inserted into each other in the intergenic spaces (SanMiguel *et al*., 1996, 1998, SanMiguel and Bennetzen, 1996). Preferential location seems to be found only for Rot 18 which shows a distribution in a cluster-like organization on stronger DAPI regions.

Although we could not explain what kind of sequences are responsible for the overall dispersion of epigenetic marks associated with heterochromatin in *Fagus sylvatica*, we disclose that repetitive sequences, have essentially a disperse distribution in Fagaceae species, including retrotransposon-like sequences in *Quercus serrata* and that some repetitive sequences can be species –specific.

# *VI. Epigenetic marks in the mature pollen of Quercus suber L. (Fagaceae)*

Sex Plant Reprod
DOI 10.1007/s00497-008-0083-y

ORIGINAL ARTICLE

# Epigenetic marks in the mature pollen of *Quercus suber* L. (Fagaceae)

**Teresa Ribeiro · Wanda Viegas · Leonor Morais-Cecílio**

Received: 12 August 2008 / Accepted: 13 September 2008


**Abstract** We have analysed the distribution of epigenetic marks for histone modifications at lysine residues H3 and H4, and DNA methylation, in the nuclei of mature pollen cells of the Angiosperm tree *Quercus suber*; a monoecious wind pollinated species with a protandrous system, and a long post-pollination period. The ultrasonic treatment developed for the isolation of pollen nuclei proved to be a fast and reliable method, preventing the interference of cell wall autofluorescence in the in situ immunolabelling assays. In contrast with previous studies on herbaceous species with short progamic phases, our results are consistent with a high level of silent (5-mC and H3K9me2) epigenetic marks on chromatin of the generative nucleus, and the prevalence of active marks (H3K9me3 and H4Kac) in the vegetative nucleus. The findings are discussed in terms of the pollination/fertilization timing strategy adopted by this plant species.

**Keywords** Epigenetic marks · Vegetative nucleus · Generative nucleus · Fagaceae · Delayed fertilization

Communicated by Thomas Dresselhaus.

T. Ribeiro (✉) · W. Viegas · L. Morais-Cecílio
Departamento de Botânica e Engenharia Biológica,
Centro de Botânica Aplicada à Agricultura (CBAA),
Instituto Superior de Agronomia, Technical University
of Lisbon, Tapada da Ajuda, 1349-017 Lisbon, Portugal
e-mail: atribeiro@isa.utl.pt

L. Morais-Cecílio
e-mail: lmorais@isa.utl.pt

## Introduction

Chemical modifications of histones and DNA methylation patterns play an important role in regulating chromatin dynamics and transcription. The male germ cells of higher plants are initiated from haploid microspores following an asymmetric division. The nucleus of the microspore cell then migrates to one of the poles, where it undergoes a mitotic division, giving rise to two structurally and functionally different cells. The smaller and elongated cell closest to the pole, generative cell, has a reduced volume of cytoplasm containing relatively few organelles, and is predestined to divide and form the two male sperm nuclei. The larger and spherical cell vegetative cell contains higher amounts of stored molecules such as starch and lipids, and has an important role in pollen grain maturation and germination, and also in pollen tube elongation. Once formed, the generative cell rapidly detaches from the pole and floats freely within the cytoplasm of the vegetative cell, resulting in a unique ‘‘cell-within-a-cell” structure in the pollen grain.

Besides different sizes and composition, the generative and vegetative cells also present distinct nuclear configurations. Soon after the asymmetric division of the microspore, the generative nucleus enters the S phase, and then passes into G2 assuming a high chromatin condensation pattern. The vegetative nucleus is arrested at G1, and displays a diffuse chromatin appearance (Bino et al. 1990; Tanaka 1997).

Chromatin structure is not only important for DNA packaging, but also plays an active role in diverse biological processes such as gene expression, DNA replication and repair and chromosome recombination and segregation (Fuchs et al. 2006). Different patterns of chromatin organization are associated with characteristic gene expression profiles and distinctive staining properties. Euchromatin represents decondensed and potentially active chromatin,

whereas heterochromatin is condensed and largely transcriptionally inert. Heterochromatin can be constitutive, when condensation is permanent throughout the cell cycle in all tissues, or facultative when chromatin compaction is developmentally regulated as a result of cellular differentiation (Craiq 2005). Euchromatin and heterochromatin are also characterized by different epigenetic marks associated with specific posttranslational histone modifications, or with distinct levels of cytosines methylated at the position five of the pyrimidine ring (5-mC). Amongst the plethora of well-documented histone modifications, the acetylation and methylation of selected lysine residues (K) of histones H3 and H4 seem to be particularly involved in the regulation of gene transcription (Richards and Elgin 2002; Horn and Peterson 2006). Genes residing in euchromatic domains are usually associated with, the hyperacetylation of several lysine residues on histones H3 and H4, high levels of H3 methylation in K4, K36, K79, and low levels of DNA methylation. Silent heterochromatic domains are characterized by hypoacetylation of H4, and histone H3 dimethylation on K9 and K27, histone H4 methylation on K20, and DNA hypermethylation (Berger 2007). Histone methylation marks are present both on hetero- and euchromatin, being specified by the position and number of methyl groups. In plants, for instance, H3K4 methylation is restricted to euchromatin and, H3K9me1,2, to heterochromatin (review in Fuchs et al. 2006).

Angiosperms display distinct fertilization processes enabling self- or cross-pollination, which in turn can be related with the period of time from pollination to fertilisation (Weterings and Russell 2004). Studies reporting epigenetic mark patterns in pollen nuclei are scarce, and have been restricted to herbaceous plants with short progamic phases (Oakeley et al. 1997; Janousek et al. 2000, 2002; Okada et al. 2006). *Quercus suber* L. is a monoecious wind-pollinated tree species with a protandrous system and a long progamic phase, taking about 3 months between pollination and the conclusion of the fertilization process (Boavida et al. 1999). In this work we characterize the distribution of several epigenetic marks in each nucleus of the bicellular mature pollen. The study of epigenetic marks in pollen nuclei of *Q. suber* is compared with those observed in *Lilium*, in order to disclose potential correlations between different pollination/fertilization processes, as well as chromatin dynamics in pollen grain nuclei.

## Materials and methods

### Pollen collection

Catkins containing flowers after anthesis were collected from several trees in Sintra, Portugal, during the flowering season in 2008. Dehiscent anthers of *Lilium* sp. grown in the greenhouse were collected from full open flowers.

Pollen of both species was isolated by shaking the anthers into a copper sieve. The isolated pollen grains were then kept at room temperature in a sealed vials containing silica gel, to avoid any effects due to humidity until used.

### Isolation of pollen nuclei by ultrasonic treatment

Isolation of the nuclei of mature pollen was carried out according to Pan et al. (2004), with modifications. Briefly, isolated pollen grains were resuspended in nuclear isolation buffer for woody plants (WPB) (Loureiro et al. 2007), consisting of 0.2 M Tris HCl, 4 mM $MgCl_2.6H_2O$, 2 mM EDTA $Na_2.2H_2O$, 86 mM NaCl, 10 mM sodium metabisulfite, 1% (w/v) PVP-10, 1% (v/v) Triton X-100, at pH 7.5. The suspension of pollen grains (0.5 ml of WPB per 0.01 $cm^3$ of pollen grains) was treated on ice with an ultrasonic device (Ultrasonic Processor, model UP 50 H, Dr. Hielsher GmbH) at 40 W, using 0.8 cycle with 80% amplitude, for 5 min, or with a low magnitude (40 W, 0.5 cycles with 40% magnitude, for 3 min). The treated solution was then filtered through miracloth (Calbiochem) and centrifuged at 500*g* for 5 min. According to the amount of pellet obtained, part of the supernatant was carefully discarded and the pellet resuspended in approximately 100 µL of the remaining volume.

### Pre-treatment of microscope slides

For enhanced nuclei adhesion, microscope slides were previously immersed in a freshly prepared solution of 2% (v/v) APTES (3-aminopropyltriethoxy-silane, Sigma) in acetone for 10 s, followed by a brief wash in acetone and water, and finally air-dried.

### Slide preparation and immunostaining of pollen nuclei

For histone immunoanalysis, nuclei suspension aliquots were mounted on APTES-coated slides, and attached to slides by the dry-ice technique (Schwarzacher and Heslop-Harrison 2000). Fixation was carried out on the slides with 200 µL of 4% (w/v) formaldehyde in 1× PBS during 20 min at room temperature, and further washed for 5 min in 1× PBS. Immunodetection of modified histones was done according to published methods (Houben et al. 2003). To avoid non-specific antibody binding, slides were incubated for 30 min in a blocking solution of 8% (w/v) BSA, 0,1% (v/v) Triton X-100 in PBS at room temperature. After washing in 1× PBS, slides were incubated for 12 h at 4°C in a humid chamber with primary antibodies in the following dilutions: 1:200 (anti-H3K9Ac, Upstate 06-942), 1:500 (anti-H3K9me2, Abcam

AB 7312, and anti-H4K5,K8,K12,K16Ac, Abcam AB 2380) and 1:1,000 (anti-H3K4me3, Abcam AB 8580) in 1× PBS supplemented with 1% BSA. The slides were further washed in 1× PBS and incubated in Cy3-conjugated anti-rabbit IgG diluted 1:100 in 1× PBS, 1% BSA for 1 h at 37°C. After final washes in 1× PBS, slides were mounted in antifade containing DAPI as DNA counterstain.

For cytosine methylation immunoanalyses the suspensions of nuclei were mixed with an equal volume of the fixative 3:1 (v/v) ethanol:glacial acetic acid mixture, placed on ice for 10 min and centrifuged 5 min at 500*g*. The supernatant was discarded and replaced with fresh 3:1 (v/v) ethanol:glacial acetic acid mixture, and this fixation and centrifugation was repeated. According to the amount of pellet obtained, a volume (approximately 100 µL) of new 3:1 fixative was adjusted and the pellet resuspended. Aliquots were dropped on to APTES-coated slides and allowed to dry. Immunolocalization of 5-methylcytosine residues was performed according to published methods (Castilho et al. 1999). Briefly, slides were blocked for 30 min in 1% (w/v) BSA in 1× PBS supplemented with 0.5% (v/v) Tween 20 (PBST) at room temperature, and after washing in 1× PBST were incubated for 1 h at 37°C in a humid chamber with a mouse monoclonal antibody raised against 5-methylcytosine (anti-5-mC, Abcam AB 10805) diluted 1:200 in blocking solution. Slides were subsequently washed in 1× PBST and then incubated in Cy3-conjugated anti-mouse IgG diluted 1:100 in 1× PBS, 1% BSA for 1 h at 37°C. Finally, slides were washed in 1× PBST and mounted in antifade containing DAPI as a DNA counterstain.

### Image capturing

Fluorescence signals were analysed under an epifluorescence microscope (Zeiss Axioscop2, Zeiss, Jena, Germany) with the appropriate filters. The images recorded with the Carl Zeiss Axiocam digital camera and Axiovision software (version 3.0.6.38), were separately captured and merged with Photoshop® (Adobe® Systems Inc., San Jose, California, USA). Nuclei area measurements were performed using the Axiovision measurement module 3.0.0.0 (Zeiss).

## Results

### Ultrasonic isolation of pollen nuclei maintains chromatin structure and integrity

A low magnitude ultrasonic treatment of fresh mature pollen in WPB buffer, and subsequent counterstaining with DAPI, revealed the binucleate structure of mature pollen of *Q. suber* (Fig. 1). Disruption of the exine and intine layers in one of the three cell wall apertures (white arrow, Fig. 1) allows DAPI penetration and the visualization of the two nuclei still enclosed in the pollen structure. The two nuclei could be recognized by their distinctive shape and chromatin organization: the vegetative nucleus has a nearly spherical shape. A large nucleolus detected through a faint DAPI staining (yellow arrowhead, Fig. 1), and an overall low intensity of DAPI, revealed a loose chromatin organization. The generative nucleus has an elongated shape, a small nucleolus (red arrowhead, Fig. 1) and a high intensity of DAPI staining revealing the highly condensed nature of its chromatin. Both nuclei also present marked differences in their nuclear areas, as seen in Fig. 1. A comparison between free nuclei, released from pollen grains (Fig 2a), and nuclei inside pollen grains (Fig. 1) shows that the major features of each of the two nuclear populations are maintained after ultrasonic treatment. The ratio between the areas of the vegetative and generative nuclei was also maintained both inside the pollen grains and in the free nuclei, allowing their reliable identification after immunodetection.

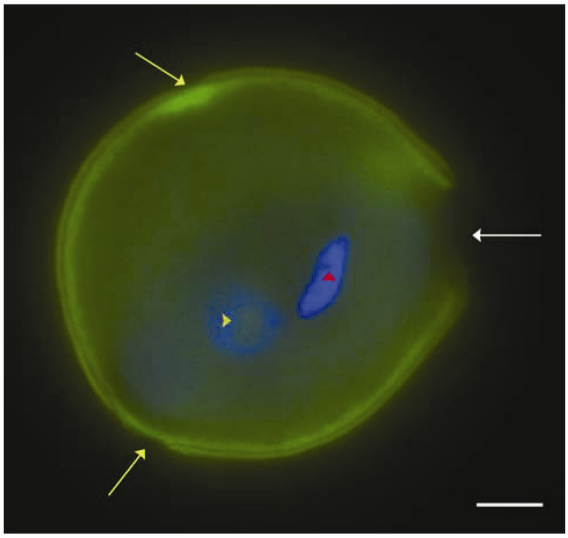


**Fig. 1** Mature pollen grain of *Quercus suber* processed with a low magnitude ultrasonic treatment and stained with DAPI. Pollen wall is visible due to autofluorescence. Pollen grains have three apertures which were differently affected by the treatment: one aperture shows complete rupture (*white arrow*), while the others still remain closed (*yellow arrows*). The two nuclei of the bicellular pollen are differently stained by DAPI (*blue signal*). The *yellow arrowhead* indicates the large nucleolus of the vegetative nucleus and the *red arrowhead*, the small nucleolus of the generative nuclei. *Bar* = 10 µm

Nuclei of *Lilium* sp. were used as a control for the validation of the nuclear isolation technique in studies of nuclear organization, and in the characterization of immunodetection patterns.

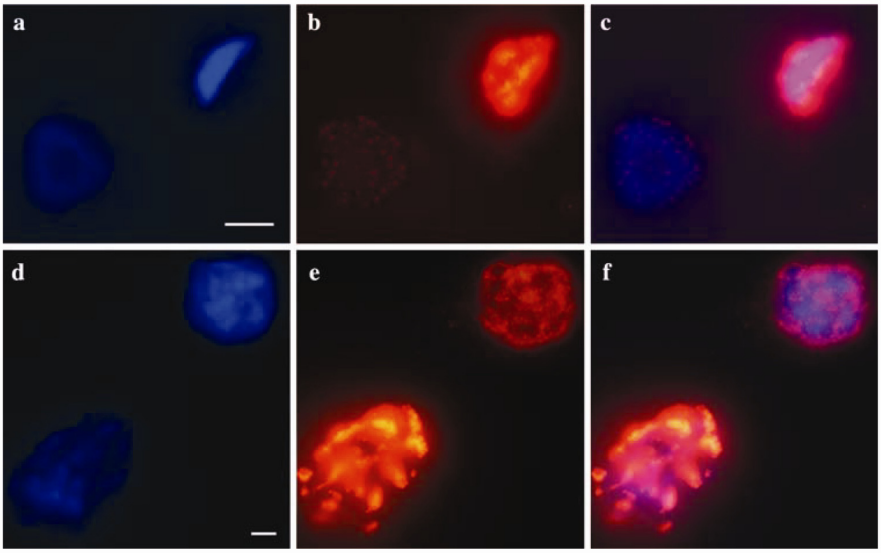



**Fig. 2** Comparative distribution of methylated cytosines (5-mC) in ultrasonic treatment- isolated vegetative and generative nuclei from *Quercus suber* (**a**, **b**, **c**), and from *Lillium* sp. pollen grains (**d**, **e**, **f**). DAPI staining of vegetative (*left*) and generative (*right*) nuclei (**a**, **d**). Immunodetection of 5-mC residues in both nuclei (**b**, **e**). Merged images of DAPI staining and immunodetection (**d**, **f**). *Bars* = 10 μm

*Lilium* isolated nuclei have statistically significant (*t*-test, $P < 0.001$) larger areas (mean values for vegetative = 1,590 ± 380 $\mu m^2$, $n = 30$ and generative = 1,190 ± 280 $\mu m^2$, $n = 30$) than *Q. suber* (mean values for vegetative 195 ± 42 $\mu m^2$, $n = 30$ and generative nuclei 73 $\mu m^2$ ± 14, $n = 30$). The differences can be attributed to the relative *C* values of cork oak ($1C = 931$ Mbp) and *Lilium* sp. Although the *Lilium* species used is unidentified, the DNA content data available for this genus ranges from $1C = 13{,}426$–46,942 Mbp (Bennett and Leitch 2004). The size proportion between vegetative and generative nuclei is more marked in *Q. suber* (mean value 0.38 ± 0.05) (Fig. 2a) than in *Lilium* (mean value 0.74 ± 0.09) (Fig. 2d).

## Distribution patterns of 5-mC residues in vegetative and generative nuclei of *Quercus suber* and *Lilium* sp.

In addition to preserving nuclear structure and integrity the isolation of pollen by ultrasound also avoids pollen wall autofluorescence, and allows antibodies' accessibility.

In *Q. suber* the labelling pattern of the anti-5-mC antibody is highly distinctive in both the vegetative and the generative nuclei (Figs. 2b, 3). The vegetative nucleus displays a very weak label, with a discrete spotty pattern (Fig. 3) only detected when the generative nucleus is overexposed (Fig. 2b). The generative nuclei in contrast are entirely labelled, showing regions with different intensities (Fig. 3). In *Lilium* sp. in contrast differences in 5-mC patterns between the two kinds of nuclei are not very marked, although the vegetative nucleus presents a stronger labelling than the generative one, (Fig. 2e) in agreement with the previous description (Janousek et al. 2000).

## Distribution patterns of histones modifications in mature pollen nuclei of *Quercus suber*

To characterize chromatin epigenetic marks in pollen grain nuclei of *Q. suber* the following antibodies for post-translation histone modifications were used: histone H3 trimethylated at lysine 4 (H3K4me3), dimethylated at lysine 9 (H3K9me2) and acetylated at lysine 9 (H3K9Ac), and histone H4 tetra-acetylated at lysines 5, 8, 12, 16 (H4tetraAc). Figure 3 shows that the distribution patterns of all the histones modifications studied are markedly different between the two nuclei, with the generative nuclei always more strongly labelled than the vegetative one (Table 1). All epigenetic marks studied are distributed throughout the generative nuclei, whereas in the vegetative ones they present a discrete spotty pattern with random distribution. In vegetative nuclei, H4tetraAc and H3K4me3 isoforms show a stronger label when compared with the other isoforms (Fig. 3; Table 1) while in generative nuclei all isoforms show a very strong label except for a weaker H3K9Ac.

The lysine residue of histone H3 at the position nine (H3K9) can be either methylated or acetylated, and both modifications present distinct patterns for each type of nucleus. In generative nuclei the level of H3 acetylation is more reduced than dimethylation (Table 1), whereas in the vegetative nuclei both H3K9 modifications have similar intensities and distribution patterns (Fig. 3; Table 1).

In summary, *Q. suber* pollen nuclei show a 5-mC pattern opposite to the one in *Lilium* sp. (Fig. 2). *Q. suber* vegetative and generative nuclei accumulate simultaneously epigenetic marks related with gene activity and gene silencing. All epigenetic marks are more represented in

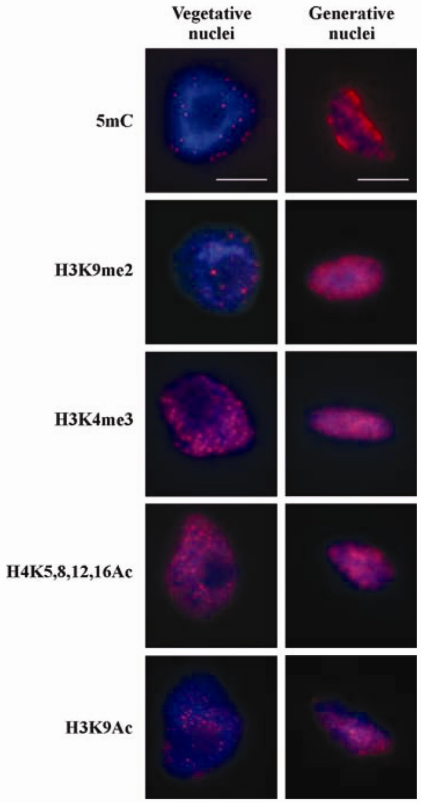



**Fig. 3** Vegetative and generative pollen grain nuclei of *Quercus suber*. Distribution patterns of epigenetic marks associated with heterochromatin (DNA methylation, 5-mC and histone H3 dimethylated at lysine 9 (H3K9me2); euchromatin (histone H3 trimethylated at lysine 4 (H3K4me3), and acetylated isoforms of histones H4 (H4K5,K8,K12,K16Ac) and H3 (H3K9Ac). The exposure time in the DAPI staining and immunodetection capturing was adjusted in both vegetative and generative nuclei in order to discriminate the labelling pattern of different epigenetic marks. *Bars* = 10 μm

generative nuclei rather than in vegetatives (Fig. 3; Table 1).

**Table 1** Comparative intensity of epigenetic marks between the vegetative and generative nuclei of *Quercus suber* mature pollen, from immunolabelling

| Epigenetic marks | *Quercus suber* nuclei | |
|---|---|---|
| | Vegetative | Generative |
| 5-mC | + | ++++ |
| H3K9me2 | + | ++++ |
| H3K4me3 | +++ | ++++ |
| H4K5,K8,K12,K16Ac | ++ | ++++ |
| H3K9Ac | + | ++ |

*5-mC* 5-methyl cytosine, *H3K9me2* histone H3 dimethylated at lysisne 9, *H3K4me3* histone H3 trimethylated at lysisne 4, *H4K5,K8,K12,K16Ac* histone H4 acetylated at lysisnes 5, 8, 12 and 16, *H3K9Ac* histone H3 acetylated at lysisne 9

Epigenetic marks label intensity visually evaluated. + Weak and scarce, ++ intermediate, +++ strong, ++++ very strong

## Discussion

An integrated study of the epigenome of *Q. suber*, addressing the pattern of several epigenetic marks associated with gene silencing (cytosine methylation, 5-mC, and histone modifications, H3K9me) and with gene activity (histone modifications, H3K4me3, H3K9Ac and H4tetra-Ac) in both nuclei of the mature pollen was performed using an ultrasonic treatment for nuclei isolation.

To our knowledge this work uses ultrasonic treatment in the isolation of nuclei for in situ immunodetection assays for the first time. Our approach shows that by avoiding the fluorescence constraints of pollen grain walls, and the presence of cytoplasm, not only can the two nuclear populations be perfectly recognized, but easy access of the antibodies to the chromatin also made possible the accurate characterization of distinct epigenetic marks.

Analysis of *Lilium* sp. pollen nuclei revealed a 5-mC distribution pattern similar to the one previously described (Janousek et al. 2000), with more intense label in the vegetative than in the generative nucleus, fully validating ultrasonic treatment for studies of chromatin organization.

The epigenetic marks studied have distinct patterns distinguishing vegetative and generative nuclei of *Q. suber*, and are different from those described in other species. Although the generative nuclei of cork oak present highly condensed chromatin they also display some features normally associated with gene activity, such as immuno signals for H4KAc and H3K4me3, as was also observed in *Lilium* (Janousek et al. 2000; Okada et al. 2006). Recent studies concerning transcriptome profiling in the generative nuclei of *Lilium longiflorum* revealed the upregulation of 356 genes correlating epigenetic and gene expression patterns (Okada et al. 2007). The level of H4 acetylation has also been related with DNA replication and/or post-replication processes, such as DNA repair, maintenance of DNA methylation, and chromatin remodelling (Fuchs et al. 2006).

Comparative analysis of epigenetic marks in both nuclei of cork oak pollen revealed stronger label associated with gene silencing (H3K9me2 and 5-mC) in generative than in vegetative nucleus, which presents a more relaxed chromatin structure. In *Lilium,* a strong label of H3K9me2 was observed in the generative nucleus (Okada et al. 2006), whereas the level of 5-mC was higher in the vegetative one (Janousek et al. 2000). In other herbaceous species studied, with either bicellular or tricellular mature pollen grains, a marked DNA methylation in the vegetative nucleus was

also observed, although associated with a more relaxed chromatin organization compared with the generative one (Oakeley et al. 1997; Janousek et al. 2002). This apparent contradiction has been interpreted on the basis that DNA cytosine hypermethylation in the vegetative nucleus reflects temporary inactivation mechanisms acting during pollen grain maturation, which is essential for their long-term survival until they reach the stigma (Janousek et al. 2002).

In *Q. suber* the vegetative nuclei present a coherent pattern of epigenetic marks with reduced amounts of chromatin silencing marks (5-mC, and H3K9me) and high levels of active chromatin marks (H3K4me3, and H4tetraAc). This pattern contrasts with the faint or nearly absent labelling of H3K4me2 (Okada et al. 2006) and H4KAc (Janousek et al. 2000) described for *Lilium* vegetative nuclei, emphasizing the differences in chromatin patterns between the herbaceous plants and the oak tree pollen grains.

We also analysed the distribution of histone H3 modifications at the lysine 9 (methylation, H3K9me2 and acetylation, H3K9Ac) in *Q. suber* pollen nuclei. The presence of the two types of H3K9 modifications was detected in both nuclei, although fainter signals were observed in vegetative ones. Lysine 9 of H3 is more methylated than acetylated in generative nuclei, whereas vegetative nuclei reveal more acetylation than methylation. While H3K9me2 is related with gene silencing, a clear correlation between transcription and H3K9Ac isoform has not yet been established in plants. It is known, however, that the presence of H3K9Ac is not cell-cycle or replication dependent, occurring in both heterochromatic and euchromatic domains of *Vicia faba* and *Hordeum vulgare*, as well as in the euchromatic domains of *Arabidopsis* (Jasencakova et al. 2000, 2001, 2003).

Altogether our analysis of the epigenetic marks on *Q. suber* pollen grains discloses patterns that are distinct from those previously described, suggesting the existence of a correlation with different pollination/fertilization strategies in angiosperms. The chromatin compaction level of the generative nuclei is also more pronounced in *Q. suber* than in *Lilium* sp., which could reflect different levels of transcription between these species, relative to differences in their pollination/fertilization processes. During the progamic stage the generative nucleus needs to maintain its integrity to guarantee the proper transference of male genetic material to the female gametophyte, and it has to divide and achieve synchrony with its female counterpart which requires a long-range signalling between male and female gametophytes. For species with very long progamic phases, such as *Q. suber*, this process is even more essential to successfully achieve fertilization. We propose therefore that the complex process underlying the required synchrony between male and female gametophytes is associated with specific patterns of high order chromatin organization.

This hypothesis considers that the low level of DNA methylation associated with histone modifications characteristic of active chromatin in the *Q. suber* vegetative nucleus, indicates a high potential for transcription that can compensate the apparent chromatin silencing of the generative nucleus.

In conclusion, our data further support the idea that the epigenetic patterns associated with chromatin organization in mature nuclei of pollen grains are not universal, but reflect the fertilization timing strategies adopted by each plant species.

**Acknowledgments** We thank Dr. Andreas Houben (Leibniz Institute of Plant Genetics and Crop Plant Research) for kindly providing antibody against H3K9Ac. We are grateful to Augusta Barão for excellent technical assistance and Xana Castilho for her involvement with the manuscript. We also wish to thank Prof Neil Jones for his critical revision of the manuscript and editing of English. TR was supported by Fundação para a Ciência e Tecnologia with grant SFRH/BD/13319/2003.

## References

Bennett M, Leitch I (2004) Angiosperm DNA c-values database, release 5.0, Dec 2004. http://www.kew.org/cval/homepage.html

Berger SL (2007) The complex language of chromatin regulation during transcription. Nature 447:407–412

Bino RJ, Tuyl JM, De Vries JN (1990) Flow cytometric determination of relative nuclear DNA contents in bicellulate and tricellulate pollen. Ann Bot 65:3–8

Boavida LC, Varela MC, Feijó JA (1999) Sexual reproduction in the cork oak (*Quercus suber* L.). I. The progamic phase. Sex Plant Reprod 11:347–353

Castilho A, Neves N, Rufini-Castiglione M, Viegas W, Heslop-Harrison JS (1999) 5-Methylcytosine distribution and genome organization in Triticale before and after treatment with 5-azacytidine. J Cell Sci 112:4397–4404

Craiq JM (2005) Heterochromatin-many flavours, common themes. Bioessays 27:17–28

Fuchs J, Demidov D, Houben A, Schubert I (2006) Chromosomal histone modification patterns—from conservation to diversity. Trends Plant Sci 11:199–208

Horn PJ, Peterson CL (2006) Heterochromatin assembly: a new twist on an old model. Chromosome Res 14:83–94

Houben A, Demidov D, Gernand D, Meister A, Leach CR, Schubert I (2003) Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. Plant J 33:967–973

Janousek B, Zluvova J, Vyskot B (2000) Histone H4 acetylation and DNA methylation dynamics during pollen development. Protoplasma 211:116–122

Janousek B, Matsunaga S, Kejnovsky E, Zluvova J, Vyskot B (2002) DNA methylation analysis of a male reproductive organ specific gene (MROS1) during pollen development. Genome 45:930–937

Jasencakova Z, Meister A, Walter J, Turner BM, Schubert I (2000) Histone H4 acetylation of euchromatin and heterochromatin is cell cycle dependent and correlated with replication rather than with transcription. Plant Cell 12:2087–2100

Jasencakova Z, Meister A, Schubert I (2001) Chromatin organization and its relation to replication and histone acetylation during the cell cycle in barley. Chromosoma 110:83–92

Jasencakova Z, Soppe WJ, Turner BM, Schubert I (2003) Histone modifications in *Arabidopsis*-high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. Plant J 33:471–480

Loureiro J, Rodriguez E, Dolezel J, Santos C (2007) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. Ann Bot 100:875–888

Oakeley J, Podesta A, Jost JP (1997) Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation. Proc Natl Acad Sci 94:11721–11725

Okada T, Singh MB, Bhalla PL (2006) Histone H3 variants in male gametic cells of lily and H3 methylation in mature pollen. Plant Mol Biol 62:503–512

Okada T, Singh MB, Bhalla PL (2007) Transcriptome profiling of *Lilium longiflorum* generative cells by cDNA microarray. Plant Cell Rep 26:1045–1052

Pan G, Zhou Y, Fowke LC, Wang H (2004) An efficient method for flow cytometric analysis of pollen and detection of 2n nuclei in *Brassica napus* pollen. Plant Cell Rep 23:196–202

Richards EJ, Elgin SC (2002) Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. Cell 108:489–500

Schwarzacher T, Heslop-Harrison P (2000) Practical in situ hybridization. BIOS Scientific Publishers, Oxford, p 203

Tanaka I (1997) Differentiation of generative and vegetative cells in angiosperm pollen. Sex Plant Reprod 10:1–7

Weterings K, Russell SD (2004) Experimental analysis of the fertilization process. Plant Cell 16:S107–S118

# *VII. Conclusions and future prospects*

Fagaceae is a family mainly composed by trees, with a small number of shrubs or treelets, enclosing several genera distributed in the Northern hemisphere from Asia to America. In European forests the most represented Fagaceae genera are *Quercus*, *Castanea* and *Fagus*.

Trees generally have slower rates of macroevolution than annual plants, what is essentially attributed to long generation time (review in Petit and Hampe, 2006). Within the three Fagaceae genera mentioned, *Quercus* is the genus with higher rate of speciation with more than 300 species in the subgenus *Quercus* (Table I.1), and does not present variability in chromosome number between species, which contradict the correlation suggested between rates of speciation and karyotype evolution, namely the variability in chromosome number (Levin and Wilson, 1976).

Although *Quercus*, *Castanea* and *Fagus* genera show the same chromosome number, 2n=24, their species present several distinct features such as the number and location of rRNA genes. *Fagus* represented by *F. sylvatica* shows four 45S rDNA *loci* while all the *Quercus* and *Castanea* species show only two *loci*. Interestingly, only one NOR seems to be active in all the species studied. Although having hundreds of species, probably due to adaptive radiation in which divergent evolution of numerous related lineages occurred within a relatively short time, acquiring for instance different growth forms like shrubs, treelets and trees, subgenus *Quercus* species seems to have "stabilized" their karyotypes, while *Castanea* genus evolved in a distinct way since the number and location of rDNA *loci* vary between species in the section *Eucastanon* with only 5 species. Although having great differences between species, the high level of homogenization of karyotypes from species of subgenus *Quercus* from different geographic regions is probably due to the spontaneous hybridization with extensive gene flow mediated by long-distance wind pollination, and high levels of introgression, although maintaining species integrity that is extremely frequent in *Quercus* (Rushton, 1993; Petit *et al*., 2003). Divergence within the genus *Castanea* is revealed by the major NOR displacement to an interstitial position and an extra 5S rDNA *locus* located adjacent to the major NOR in the Chinese chestnut *C. mollissima*, and absent in the Japanese *C. crenata* and the European *C. sativa.* The migration of *Castanea* from Japan to China and then to Europe (Lang *et al*., 2007) had an effect on karyotype structure in the Chinese chestnut *C. mollissima.* The fixation of chromosomal rearrangements

should have occurred after the speciation of *C. mollissima*, since *C. sativa* maintained the same number and disposition of rDNA *loci* as *C. crenata*. And probably have occurred before the diversification between *C. mollissima* and the other two Chinese species, *C. seguinii and C. henryi* which seems to have occurred relatively recent in *Castanea* evolution (Lang *et al*., 2007). A possible explanation is the founder effect considering that a population of *C. mollissima* could have experienced bottleneck selection reducing the population. Since mutations become more rapidly fixed in a small population by genetic drift and because *Castanea* is a monoecious genus with also some bisexual catkins, selfing lifestyle could have been favoured leading to inbreeding and consequently to homozigoty. The two detected chromosome rearrangements drastically affected nucleolar chromosome structure leading to the lost of the NOR major pseudosatellite structure in *C. mollissima*.

The number and location of 45S and 5S rDNA *loci* in subgenus *Quercus* spp. (Zoldoš *et al*., 1999, and our data) and in two *Castanea* spp. as also been detected in the majority of the studied species of Fagaceae belonging to *Lithocarpus*, *Quercus* subgenus *Cyclobalanopsis* and *Castanopsis* genera (Chokchaichamnankit *et al*., 2008), contrasting with the higher number of 45S and 5S rDNA *loci* observed in *Fagus sylvatica*. Variations in the distribution of rDNA *loci* have direct phylogenetic implications, since *taxa* proximity has been correlated in other species of trees with FISH mapping (Hizume *et al*., 2002; Liu *et al*., 2003). The rDNA mapping and the morphology of the bigger chromosome one is in accordance with an early divergence of *Fagus* and the closeness of *Quercus* and *Castanea* also suggested by phylogenetic analyses based on molecular data (Manos *et al.,* 2001; Li *et al.,* 2004). However our comparisons of whole karyotypes through morphometric analysis and asymmetry indices suggest that *Castanea* is clearly divergent from *Quercus*, what can be attributed to the differential amplification of some repetitive sequences.

Therefore we support the early divergence of *Fagus* and the closeness of *Quercus* and *Castanea* by rDNA mapping, differential presence of the pseudosatellite, and the high homology of faster evolving repetitive sequences such a retrotransposon-like sequence (Rot 10) or the 45S ETS regions between *Quercus* and *Castanea* genera. Although trees are known for their slow macroevolution, it is still strange such high homology between fast evolving sequences found in this work between two genera that have diverged 60 mya (Manos and Steele, 1997; Manos *et al*., 2001).

Analysis of Fagaceae IGS was also very important revealing besides the canonical head to tail organization of rDNA repeats, internal rearrangements that might be in the constant process of concerted evolution. The ETS regions flanked by palindromic partial 18S sequences showed distinct lengths being the *F. sylvatica* the smaller one indicating either a deletion in the *Fagus* genome occurring after the divergence of *Castanea* and *Quercus* from *Fagus* or an amplification in the *Quercus* and *Castanea* genomes which occurred just after their speciation process.

In addition to the genetic involvement we cannot ignore the contribution of epigenetics to genome evolution. Indeed, diversity between genomes results from the interplay between DNA sequence evolution and epigenetic evolution (Grant-Downton and Dickinson, 2006). Distribution of repetitive sequences in small and large genomes is thought to be correlated with the different types of interphase chromatin organization detected in these genomes. However different *Quercus* species with similar genome sizes (Zoldoš *et al*., 1998) show distinct patterns of heterochromatin interphase organization revealed by the presence or absence of interphase chromocenters. Our observations clearly indicate that neither genome size nor even the amount of DNA per chromosome are directly correlated with interphase chromatin organization. The absence of a correlation between DNA contents and distribution patterns of heterochromatin can also be observed in *Ornithogalum longebracteanum* (7448 Mb/1C), a large genome species, (~276Mb/chromosome) which presents an interphase chromatin distribution similar to the one of Arabidopsis, a small genome species with 155Mb/1C and ~31Mb/chromosome (Bennett and Leitch 2004).

The distribution of epigenetic marks in genomes of *Quercus* and *Castanea* species without chromocenters are in accordance to the ones detected in large or medium size genomes *i.e*. distributed all over the genome (Houben *et al*., 2003; Castilho *et al*., 1999). On the other hand, the disposition of silent associated epigenetic marks in *F. sylvatica* is not restricted to the chromocenters, revealing therefore a different distribution of heterochromatin-associated epigenetic marks in nuclei of plants with small genomes. Furthermore, epigenetic marks of rDNA domains clearly revealed silent and active *loci* in *F. sylvatica*, while in *Quercus* and *Castanea* species only the silent domains are undoubtedly detected being the pesudosatellite in *C. sativa* the most distinctive genome domain revealing a highly methylated status. These epigenetic patterns can be related with the different types of interphase chromatin organization present in species of the three genera.

Nuclear organization is affected by distribution of repetitive sequences and also by cell cycle phase, development and cell differentiation. *Quercus suber* mature pollen grains present besides patterns similar to the ones described in other species (Janousek *et al*., 2000; Okada *et al*., 2006) specific epigenetic patterns disclosed in this work. For instance a strong presence of epigenetic marks related with gene silencing as well as with gene transcription in the generative nuclei besides its heterochromatic appearance was observed. The long held view that generative nuclei were transcriptional silent was recently set apart (Okada *et al*., 2007); in fact it is understandable that proteins necessary for division have to be transcribed in these nuclei. On the other hand the vegetative nuclei of *Q. suber* show comparatively high presence of epigenetic marks associated with gene transcription and reduced incidence of epigenetic marks associated with gene silencing, especially 5-mC. These findings are not in agreement with the normally accepted transcriptionally quiescent state of these nuclei until the formation of the pollination tube in herbaceous species (Oakeley *et al*., 1997; Janousek *et al*., 2000; Janousek *et al*., 2002). The difference in methylation level between generative and vegetative nuclei detected in herbaceous plants (Oakeley *et al*., 1997; Janousek *et al*., 2000; Janousek *et al*., 2002) and *Q. suber* which shows delayed fertilization suggests that the program of pollen cells differentiation although sharing most of the mechanisms is not universal and may depend upon the fertilization strategy adopted by the species.

Abnormalities during microsporogenesis can lead to 2n pollen which is the most common process attributed to the formation of triploids trees (Butorina, 1993; Zhang *et al.* 2007). Nevertheless, we could not detect 2n pollen in *Q. suber* trees that could justify the *Q. suber* triploid individual characterized in this study. It has been suggested that the presence of triploids could be an evolutionary bridge to the occurrence of tetraploids (Baack, 2005); however no tetraploids have ever been detected in the subgenus *Quercus*. Polyploidy has been found in one Fagaceae species, *Trigonobalanus verticillata* 2n=6x=42 (Hou, 1971) considered as the others *Trigonobalanus* spp. *sensu lato* (*Colombobalanus excelsa* and *Formanodendron doichangensis*) relic species (Manos *et al*., 2001; Kamiya *et al*., 2002). On the other hand, paleopolyploidy is an increasingly known event in plant evolution (review Eckardt, 2004), and recently it was found that a large-scale gene-duplication event had occurred in the ancestor of poplar, or at least very early in the evolution of the *Populus* genus (Sterck *et al*., 2005). The basic number x=7 found in *T. verticillata* (Hou, 1971), *F. doichangensis* (*T. doichangensis*) (2n=14) (Chen *et al*., 2007) and in one tree of *Q. lenticellatus* (2n=14)

(Chokchaichmnankit *et al*., 2008) belonging to the subgenus *Cyclobalanopsis,* as well as the basic number x= 6 in all other Fagaceae species (2n=24) may indicate a paleopolyploidy event in the Fagaceae family. According to Stebbins (1971) basic numbers of modern woody genera were derived by ancient polyploidy, and the original basic numbers of angiosperms, both woody and herbaceous, were x = 6 and x = 7.

Although the family Fagaceae dominates forests in the temperate regions of the Northern Hemisphere its centre of diversity was found in tropical South-East Asia, particularly at the generic level (Manos *et al.,* 2001). Briggs and Walters (1997) hypothesized that temperate woody groups have, in general, been derived from tropical ancestors. Moreover, it is in the tropical South-East Asia that it is believed to be maintained the most primitive forms of *Castanopsis* and *Lithocarpus* as well as *Quercus*, subgenus *Cyclobalanopsis* (Soepadmo, 1972), besides the species *T. verticillata, F. doichangensis* (Chen *et al*., 2007) and *Q. lenticellatus* (Chokchaichmnankit *et al*., 2008)

The modern karyotypes of Fagaceae may have derived from ancestors, with the basic number x=7 which by dysploidy became x=6, following by allopolyploid or autopolyploid events that ended in the actual 2n=2x=24. High level of polyploidy such as the one present in *Trigonobalanus verticillata* may have not give any karyotype advantage and therefore became an evolutionary "dead end". The double rDNA loci found in the basal genus *Fagus* could be a remnant of that ancestral duplication event. However, duplicated segments could not be detected in the several genetic linkage maps of *Fagus*, *Quercus*, and *Castanea* (Barreneche *et al*., 1998, 2004; Casasoli *et al*., 2001, 2006; Scalfi *et al*., 2004), as it has been revealed in genetic linkage maps of the diploid *Brassica oleraceae* (Bouhon *et al*., 1996; Slocum *et al*., 1990). This difference can be due to the time polyploidy events took place in Fagaceae and in Brassicas.

From this work several questions arose and new lines of investigations are pointed out, namely:

- To unravel the karyotype evolution in the Fagaceae family, since several groups have not yet been characterized at the chromosome level. Furthermore in species such as *F. doichangensis* with 14 chromosomes, the hexaploid *T. verticillata*, and species from the genus *Chrysolepis* rDNA *loci* mapping will contribute to establish chromosome evolutionary story of the family;

- While it is difficult to have a Fagaceae genome sequenced, a refine study analyzing unigene sets of ESTs and searching each set for paralogs or duplicated sequences can be implemented (Sterck, 2005, Blank and Wolfe, 2004) allowing the potential detection of an ancient polyploid origin of the Fagaceae family

- The distinct epigenetic patterns detected in generative and vegetative mature pollen nuclei in *Q. suber* implying gene transcription at the vegetative nuclei, would certainly be improved through transcriptomic analysis in order to better understand the plant strategies correlated with the delayed fertilization process.

# *VIII. References*

Abranches R., Beven A.F., Aragón-Alcaide L., Shaw P.J. 1998. Transcription sites are not correlated with chromosome territories in wheat nuclei. *J Cell Biol* **143**: 5-12.

Adams K.L., Cronn R., Percifield R., Wendel J.F.2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proc Natl Acad Sci USA* **100**: 4649-4654.

Adams K.L., Percifield R., Wendel J.F.2004. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics* **168**: 2217-2226.

Agius F., Kapoor A., Zhu J.K. 2006. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci USA* **103**: 11796-11801.

Ahmad K., Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**: 1191-1200.

Allard R.W., Saghai-Maroof M.A., Zhang Q., Jorgensen R.A. 1990. Genetic and molecular organization of ribosomal DNA (rDNA) variants in wild and cultivated barley. *Genetics* **126**: 743-751

Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res* **25**: 389-402.

Anamthawat-Jónsson K., Thórsson A.T. 2003. Natural hybridisation in birch: Triploid hybrids between *Betula nana* and *B. pubescens*. *Plant Cell Tiss Org Cult* **75**:99-107.

Arnheim N., Krystal M., Schmickel R., Wilson G., Ryder O., Zimmer E. 1980. Molecular evidence for genetic exchanges among ribosomal genes on non-homologous chromosomes in man and apes. *Proc Natl Acad Sci USA* **77**: 7323–7327.

Arumuganathan K., Earle E.D. 1991. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* **9**: 208-218.

Avramova Z., Tikhonov A., Chen M., Bennetzen J.L. 1998. Matrix attachment regions and structural colinearity in the genomes of two grass species. *Nucleic Acids Res* **26**: 761–767.

Baack E. 2005. To succeed globally, disperse locally: effects of local pollen and seed dispersal on tetraploid establishment. *Heredity* **94**: 538–546.

Bacilieri R., Ducousso A., Petit R.J., Kremer A. 1996. Mating system and asymmetric hybridization in a mixed stand of European oaks. *Evolution* **50**: 900-908.

Barcaccia G., Meneghetti S., Albertini E., Triest L., Lucchin M. 2003. Linkage mapping in tetraploid willows: Segregation of molecular markers and estimation of linkage phases support an allotetraploid structure for *Salix alba* x *Salix fragilis* interspecific hybrids. *Heredity* **90**:169-180.

Barow M., Meister A. 2002. Lack of correlation between AT frequency and genome size in higher plants and the effect of nonrandomness of base sequences on dye binding. *Cytometry* **47**: 1–7.

Barral G., Poggio L., Giberti G. 1995. Chromosome numbers and DNA content from Ilex argentina (Aquifoliaceae). *Bol Soc Argent Bot* **30**:243-248.

Barreneche T., Bodenes C., Lexer C., Trontin J.F., Fluch S. *et al.*1998 A genetic linkage map of *Quercus robur* L. (peduculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. *Theor Appl Genet* **97**: 1090–1103.

Barreneche T., Casasoli M., Russell K., Akkak A., Meddour H. *et al*. 2004. Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). *Theor Appl Genet* **108**: 558–566.

Bartee L., Malagnac F., Bender J. 2001. Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev* **15**: 1753-1758.

Becker P.B., Horz W. 2002. ATP-dependent nucleosome remodeling. *Annu Rev Biochem* **71**:247-273.

Bellarosa R., Delre V., Schirone B., Maggini F. 1990. Ribosomal RNA genes in *Quercus* spp. (fagaceae). *Plant Syst Evol* **172**: 127-139.

Belmont A.S., Dietzel S., Nye A.C., Strukov Y.G., Tumbar T. 1999. Large-scale chromatin structure and function. *Curr Opin Cell Biol* **11**: 307-311.

Benevolenskaya E.V. 2007. Histone H3K4 demethylases are essential in development and differentiation. *Biochem Cell Biol* **85**: 435-443.

Bennett M.D., Smith J.B. 1976. Nuclear DNA amounts in angiosperms. *Philos T R Soc B -Biol Sci* **274**: 227-274.

Bennett M., Smith J., Heslop-Harrison J. 1982. Nuclear DNA amounts in angiosperms. *P Roy Soc Lond B- Bio Sci* **216**:179-199.

Bennett M.D., Smith J.B. 1991. Nuclear DNA amounts in Angiosperms. *Philos T R Soc B* **334**: 309-345.

Bennett M.D., Leitch I.J., Price H.J., Johnston J.S. 2003. Comparisons with Caenorhabditis (~100 Mb) and Drosophila (~ 175 Mb) using flow cytometry show genome size in Arabidopsis to be ~157 Mb and thus ~25 % larger than the Arabidopsis Genome Initiative estimate of ~125 MB. *Ann Bot* **91**: 547-557.

Bennetzen J.L. 2000a. Transposable element contributions to plant gene and genome evolution. *Plant Mol Biol* **42**: 251–269.

Bennetzen J.L. 2000b. The many hues of plant heterochromatin. *Genome Biol* **1**: reviews 107.1–107.4.

Bennetzen J.L. 2002. Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* **115**: 29–36.

Bennetzen J.L. 2005. Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr Opin Genet Dev* **15**: 621–627.

Berger S.L. 2007. The complex language of chromatin regulation during transcription. *Nature* **447**: 407-412.

Besendorfer V., Zoldoš V., Littvay T., Papes D. 1995. Karyological Characteristics of the Common Oak (*Quercus robur* L.) IUFRO XX World Congres.

Bino R.J., Tuyl J.M., De Vries J.N. 1990. Flow cytometric determination of relative nuclear DNA contents in bicellulate and tricellulate pollen. *Ann Bot* **65**: 3–8.

Bird A.P., Wolffe A.P. 1999. Methylation-Induced Repression - Belts, Braces, and Chromatin. *Cell* **99**: 451–454.

Birchler J.A., Riddle N.C., Auger D.L., Veitia R.A.2005. Dosage balance in gene regulation: biological implications. *Trends Genet* **21**: 219-226.

Blakesley D., Allen A., Pellny T.K., Roberts A.V. 2002. Natural and induced polyploidy in *Acacia dealbata* Link. and *Acacia mangium* Wild. *Ann Bot* **90**: 391–398.

Blanc G., Wolfe K.H. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* **16**: 1667–1678.

Blanton J., Gaszner M., Schedl P. 2003. Protein:protein interactions and the pairing of boundary elements in vivo. *Gene Dev* **17**: 664-675.

Boavida L.C., Silva J.P., Feijó J.A. 2001 Sexual reproduction in the cork oak (*Quercus suber* L).II. Crossing intra- and interspecific barriers. *Sex Plant Reprod* **14**:143–152.

Bobola M., Smith D., Klien A. 1992. Five major nuclear ribosomal repeats represent a large and variable fraction of the genomic DNA of *Picea rubens* and *P. mariana*. *Mol Biol Evol* **9**: 125—137.

Bode J., Stengert-Iber M., Kay V., Schlake T., Dietz-Pfeilstetter A. 1996. Scaffold/matrix-attached regions: Topological switches with multiple regulatory functions. *Crit Rev Eukaryot Gene Expr* **6**: 115–138.

Bohuon E.J.R., Keith D.J., Parkin I.A.P., Sharpe A.G., Lydiate D.J. 1996. Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. *Theor Appl Genet* **93**: 833–839.

Borgardt S.J., Nixon K.C. 2003. A comparative flower and fruit anatomical study of *Quercus acutissima*, a biennial-fruiting oak from the *Cerris* group (Fagaceae). *Am J Bot* **90**:1567-1584.

Borzan Z., Schlarbaum S. 2004. Cytogenetics of forest tree species. In: *Encyclopedia of Forest Science*. (Burley J., Evans J., Youngquist J.A. eds) Elsevier, Oxford, pp 204–214.

Brandes A., Heslop-Harrison J.S., Kamm A., Kubis S., Doudrick R.L. et al. 1997. Comparative analysis of the chromosomal and genomic organization of Ty1-copia-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Mol Biol* **33**: 11–21.

Briggs D., Walters S.M. 1997. Plant Variation and Evolution, 3rd edition. Cambridge University Press, Cambridge

Buck S.W., Sandmeier J.J., Smith J.S. 2002. RNA polymerase I propagates unidirectional spreading of rDNA silent chromatin. *Cell* **111**: 1003–1014.

Burda R.I., Shchepotiev F.L. 1973. Spontaneous polyploidy in seedlings of multi-seeded acorns of *Quercus robur* L. *Cytol Genet* **7**: 140–143.

Burger W.C. 1975. The species concept in *Quercus*. *Taxon* **24**: 45–50.

Butaye K.M., Goderis I.J., Wouters P.F., Pues J.M., Delaure S.L., Broekaert W.F., Depicker A., Cammue B.P., De Bolle M.F. 2004. Stable high-level transgene expression in Arabidopsis thaliana using gene silencing mutants and matrix attachment regions. *Plant J* **39**: 440–449.

Butorina A.K. 1993. Cytogenetic study of diploid and spontaneous triploid oaks, *Quercus robur* L. *Ann Sci For* **50**: 144–150.

Caburet S., Conti C., Schurra C., Lebofsky R., Edelstein S.J., Bensimon A. 2005. Human ribosomal RNA gene arrays display a broad range of palindromic structures. *Genome Res* **15**: 1079–1085.

Camus A. 1954. Les Chenes. Monographie du genre *Quercus*. Encyclopedie economique de sylviculture, Vol. VI, VII, VIII. Lechevalier, Paris.

Cao X., Jacobsen S.E. 2002. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci USA* **4**: 16491-16498.

Capelo J., Catry F. 2007. Biologia, ecologia e distribuição da azinheira. In: *Árvores e florestas de Portugal. Os Montados. Muito para além das árvores*. (Silva J.S. ed) Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 119-130.

Caperta A., Neves N, Morais-Cecílio L., Malhó R., Viegas W. 2002. Genome restructuring in rye affects the expression, organization and disposition of homologous rDNA loci. *J Cell Sci* **115**: 2839–2846.

Caperta A.D., Neves N., Viegas W., Pikaard C.S., Preuss S. 2007. Relationships between transcription, silver staining, and chromatin organization of nucleolar organizers in *Secale cereale*. *Protoplasma* **232**: 55-59.

Castiglione R.M., Venora G., Ravalli C., Stoilov L., Gecheff K., Cremonini R. 2008. DNA methylation and chromosomal rearrangements in reconstructed karyotypes of *Hordeum vulgare* L. *Protoplasma* **232**: 215–222.

Carmo-Fonseca M., Mendes-Soares L., Campos I. 2000. To be or not to be in the nucleolus. *Nat Cell Biol* **2**: 107–112.

Carvalho J.M. 2007. A caça e os montados de azinho. In: *Árvores e florestas de Portugal. Os Montados. Muito para além das árvores*. (Silva J.S. ed). Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 161-176.

Casasoli M., Mattioni C., Cherubini M., Villani F. 2001. A genetic linkage map of European chestnut (*Castanea sativa* Mill.) based on RAPD, ISSR and isozyme markers. *Theor Appl Genet* **102**: 1190–1199.

Casasoli M., Derory J., Morera-Dutrey C., Akkak A., Brendel O., Porth I. Guehl J.M. et al. 2006. Comparison of quantitative trait loci for adaptative traits between oak and chestnut based on expressed sequence tag consensus map. *Genetics* **172**: 533–546.

Castilho A., Neves N., Rufini-Castiglione M., Viegas W., Heslop-Harrison J.S. 1999. 5-Methylcytosine distribution and genome organization in Triticale before and after treatment with 5-azacytidine. *J Cell Sci* **112**: 4397–4404.

Castro E.B., González M.A.C., Tenorio M.C., Bombin R.E., Antón M.G., Fuster M.G., Manzaneque A.G., Manzaneque F.G., Saiz J.C.M., Juaristi C.M., Pajares P.R., Ollero H.S. 1998. In: *Los Bosques Ibéricos. Una Interpretación Geobotánica* (Tenorio, M.C., Juaristi, C.M., Ollero, H.S. eds.) Editorial Planeta, España, 597 pp.

Chalupa V. 1986. Fagus sylavatica L. (European Beech). In: *Biotechnology in Agriculture and Forestry. Trees IV* (Bajal Y.P.S. ed) Springer, pp 138-140.

Chan S.W., Zilberman D., Xie Z., Johansen L.K., Carrington J.C., Jacobsen S.E. 2004. RNA silencing genes control de novo DNA methylation. *Science* **303**: 1336.

Charlesworth B., Sniegowski P., Stephan W. 1994 The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215–220.

Chen B.F., Heneen W.K. 1995. Satellited chromosomes, nucleolus organizer regions and nucleoli of *Brassica campestris* L., *B. nigra* (L.) Koch, and *Sinapsis arvensis* L. *Hereditas* **122**: 113–118.

Chen B.F., Heneen W.K., Pedersen C. 1995. Ribosomal RNA gene loci and their nucleolar activity in *Brassica alboglabra* Bailey. *Hereditas* **123**: 169–173.

Chen Z., Pikaard C. 1997a. Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance. *Gene Dev* **11**: 2124-2136.

Chen Z.J., Pikaard C.S. 1997b. Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/ silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. *Proc Natl Acad Sci USA* **94**: 3442– 3447.

Chen D., Dundr M., Wang C., Leung A., Lamond A., Misteli T., Huang S. 2005. Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J Cell Biol* **168**: 41-54.

Chen Z.J., Tian L. 2007. Roles of dynamic and reversible histone acetylation in plant development and polyploidy. *Biochim Biophys Acta* **1769**: 295–307.

Chen G., Sun W.-B., Han C.-Y., Coombes A. 2007. Karyomorphology of the endangered *Trigonobalanus doichangensis* (A. Camus) Forman (Fagaceae) and its taxonomic and biogeographical implications. *Bot J Linn Soc* **154**:321–330.

Chokchaichmnankit P., Anamthawat-Jonsson K., Chulalaksananukul W. 2008. Chromosomal Mapping of 18S-25S and 5S Ribosomal Genes on 15 Species of Fagaceae from Northern Thailand. *Silvae Genetica* **57**: 5-13.

Clausen K., Kung F., Bey C., Daniels R.A. 1982. Variation in white ash. *Silvae Genet* **30**:93-97.

Copenhaver G.P., Pikaard C.S. 1996a. RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of Arabidopsis thaliana adjoin the telomeres on chromosomes 2 and 4. *Plant J* **9**: 259-272.

Copenhaver G.P., Pikaard C.S. 1996b. Two-dimensional RFLP analyses reveal megabase-sized clusters of rRNA gene variants in *Arabidopsis thaliana*, suggesting local spreading of variants as the mode for gene homogenization during concerted evolution. *Plant J* **9**:273-282.

Cortizo E.V., Madriñán M.L.V., Madriñán F.J.V.1996. El Castaño. León: Edilesa

Costa A., Pereira H. 2007. Montados e sobreirais: uma espécie, duas perspectivas. In: *Árvores e florestas de Portugal. Os Montados. Muito para além das árvores* (Silva J.S. ed) Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 17-37.

Cowan C.R., Carlton P.M., Cande W.Z. 2001. The polar arragement of telomeres in interphase and meiosis. Rabl organisation and the bouquet. *Plant Physiol* **125**: 532-538.

Craig J.M. 2005. Heterochromatin-many flavours, common themes. *Bioessays* **27**: 17-28.

Cremer M., von Hase J., Volm T., Brero A., Kreth G., Walter J., Fischer C., Solovei I., Cremer C., Cremer T. 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res* **9**: 541-567.

Curtu A.L., Gailing O., Finkeldey R. 2007. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. *BMC Evol Biol* **7**:218

Dane F., Lang P., Huang H., Fu Y. 2003. Intercontinental genetic divergence of *Castanea* species in eastern Asia and eastern North America. *Heredity* **91**: 314–321.

Darlington C.D., Wylie A.P. 1995. Chromosome atlas of flowering plants. George Allen and Unwin Ltd., London. 519 pp.

Dechyeva D., Schmidt T. 2006. Molecular organization of terminal repetitive DNA in Beta species. *Chromosome Res* **14**: 881-897.

D'Emerico S., Pacciola C., Tommasi F. 2000. Contribution to the karyomorphology of some species of the genus *Quercus*. *Silvae Genetica* **49**: 243–245.

Denk T. 2003. Phylogeny of *Fagus* L. (Fagaceae) based on morphological data. *Plant Syst Evol* **240**: 55–81.

Desel C., Jansen R., Dedong G., Schmidt T. 2002. Painting of parental chromatin in Beta hybrids by multi-colour fluorescent in situ hybridization. *Ann Bot* **89**: 171-181.

Dong F., Jiang J. 1998. Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. *Chromosome Res* **6**: 551-558.

Dundr M., Meier U.T., Lewis N., Rekosh D., Hammarskjöld M-L., Olson M.O.J. 1997. A class of nonribosomal nucleolar components is located in chromosome periphery and in nucleolus-derived foci during anaphase and telophase. *Chromosoma* **105**: 407–417.

Dundr M., Olson M.O.J. 1998. Partially processed pre-rRNA is preserved in association with processing components in nucleolus derived foci during mitosis. *Mol Biol Cell* **9**: 2407–2422.

Dundr M., Misteli T., Olson M.O.J. 2000. The dynamics of postmitotic reassembly of the nucleolus. *J Cell Biol* **150**: 433–446.

Dvorák J., Zhang H.-B., Kota R.S., Lassner M. 1989. Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* **32**: 1003–1016.

Dzialuk A., Chybicki I., Welc M., Sliwinska E., Burczyk J. 2007. Presence of Triploids among Oak Species. *Ann Bot* **99**:959-964.

Eckardt N.A. 2004. Two Genomes Are Better Than One: Widespread Paleopolyploidy in Plants and Evolutionary Effects. *Plant Cell* **16**: 1647-1649.

Eichler E.E., Sankoff D. 2003. Structural dynamics of eukaryotic chromosome evolution. *Science* **301**: 793-797.

Elder J.F. jr., Turner B.J. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *Q Rev Biol* **70**: 297-320.

Fang Y., Spector D.L. 2005. Centromere positioning and dynamics in living Arabidopsis Plants. *Mol Biol Cell* **16**: 5710–5718.

Favre J.M., Brown S. 1996. A flow cytometric evaluation of the nuclear DNA content and GC percent in genomes of European oak species. *Ann Sci For* **53**: 915-917.

Fernandez-Capetillo O., Mahadevaiah S.K., Celeste A., Romanienko P.J., Camerini-Otero R.D., Bonner W.M., Manova K., Burgoyne P., Nussenzweig A. 2003. H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. *Dev Cell* **4**: 497–508.

Finnegan E.J., Peacock W.J., Dennis E.S. 1996. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc Natl Acad Sci USA* **93**: 8449–8454.

Finnegan E.J., Genger R.K., Peacock W.J., Dennis E.S. 1998. DNA methylation in plants. *Annu Rev* Plant *Physiol Plant Mol Biol* **49**: 223–247.

Finnegan E.J., Peacock W.J., Dennis E.S. 2000. DNA methylation, a key regulator of plant development and other processes. *Curr Opin Genet Dev* **10**: 217–223.

Finnegan E.J., Kovac K.A., Jaligot E.S., Sheldon C.C., Peacock W.J., Dennis E.S. 2005. The down regulation of FLOWERING LOCUS C (FLC) in plants with low levels of DNA methylation and by vernalization occurs by distinct mechanisms. *Plant J* **44**: 420-442.

Fisher A., Hofmann I., Naumman K., Reuter G. 2006. Heterochromatin proteins and the control of heterochromatic gene silencing in Arabidopsis. *J Plant Physiol* **163**: 358-368.

Flavell R.B., Bennett M.D., Smith J.B., Smith D.B. 1974. Genome size and proportion of repeated nucleotide DNA sequence in plants. *Biochem Genet* **12**: 257–269.

Flavell R.B. 1986a. Repetitive DNA and chromosome evolution in plants. *Philos T R Soc B* **312**: 227–242.

Flavell R.B., O'Dell M., Thompson W.F. 1988. Regulation of cytosine methylation in ribosomal DNA and nucleolus organizer expression in wheat. *J Mol Biol* **204**: 523–534.

Forman L.L. 1964. *Trigonobalanus* a new genus of Fagaceae with notes on the classification of the family. *Kew Bull* **17**:381–396.

Fransz P., de Jong J.H., Lysak M., Castiglione M.R., Schubert I. 2002. Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. *Proc Natl Acad Sci USA* **99**: 14584-14589.

Fuchs J., Demidov D., Houben A., Schubert I. 2006. Chromosomal histone modification patterns – from conservation to diversity. *Trends Plant Sci* **11**: 199-208.

Fujisawa M., Yamagata H., Kamiya K., Nakamura M., Saji S., Kanamori H. *et al*. 2006. Sequence comparison of distal and proximal ribosomal DNA arrays in rice (*Oryza sativa* L.) chromosome 9S and analysis of their flanking regions. *Theor Appl Genet* **113**: 419–428.

Fukui K., Nakayama S. 1996. *Plant chromosomes: laboratory methods*. CRC Press, Boca Raton, Florida.

Gall J.G. 1981. Chromosome structure and the C-value paradox . *J Cell Biol* **91**: 3s-14s.

Gall J.G., Pardue M.L. 1969. Fomation and detection RNA-DNA hybrid molecules in cytological preparations. *Proc Natl Acad Sci USA* **63**: 378-383.

Gallois A., Burrus M., Brown S. 1999. Evaluation of the nuclear DNA content and GC percent in four varieties of *Fagus sylvatica* L. *Ann Forest Sci* **56**: 615- 618.

Gary J.D., Clarke S. 1998. RNA and protein interactions modulated by protein arginine methylation. *Prog Nucleic Acid Res Mol Biol* **61**: 65–131.

Gehring M., Huh J.H., Hsieh T.F., Penterman P., Choi Y., Harada J.J. et al. 2006. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation, *Cell* **124**: 495–506.

Gehring M., Henikoff S. 2007. DNA methylation dynamics in plant genomes. *Biochem Bioph Acta* **1769**: 276–286.

Gerbi S.A. 1986. The evolution of eukaryotic ribosomal DNA. *Biosystems* **19**: 247-258.

Gerlach W.L., Bedbrook J.R. 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res* **7**: 1869-1885.

Gerlach W.L., Dyer T.A. 1980. Sequence organization of the repeating units in the nucleus of wheat which contain 5S-rRNA genes. *Nucleic Acids Res* **8**:4851-4865.

Gilbert D.M. 2002. Replication timing and transcriptional control: beyond cause and effect. *Curr Opin Cell Biol* **14**: 377-383.

Gomes-Laranjo J., Peixoto F., Sang W.F.,Torres-Pereira J. 2005. Study of the temperature effect in three chestnut (Castanea sativa Mill.) cultivars' behaviour. *J Plant Physiol* **163**: 945-955.

González-Melendi P., Wells B., Beven A., Shaw P. 2001. Single ribosomal transcription units are linear, compacted Christmas trees in plant nucleoli. *Plant J* **27**: 223-233.

Goodpasture C., Bloom S.E. 1975. Visualization of nucleolar organizer regions in mammalian chromosomes using silver staining. *Chromosoma* **53**: 37–50.

Govaerts R., Frodin D.G. 1998. World checklist and bibliography of Fagales (Betulaceae, Corylaceae, Fagaceae and Ticodendraceae). Kew, UK: Royal Botanic Gardens.

Grant-Downton R.T., Dickinson H.G. 2006. Epigentics and its implications for plant biology 2. The 'epigenetic epiphany': Epigenetics, evolution and beyond. *Ann Bot* **97:** 11–27.

Grewal S.I., Elgin S.C. 2002. Heterochromatin: new possibilities for the inheritance of structure. *Curr Opin Genet Dev* **12**:178-187.

Grewal S.I., Jia S. 2007. Heterochromatin revisited. *Nat Rev Genet* **8**: 35-46.

Gruenbaum Y., Naveh-Many T., Cedar H., Bazin A. 1981. Sequence specifity of methylation in higher plant DNA. *Nature* **292**: 860-862.

Gruendler P., Unfried I., Pascher K., Schweizer D. 1991. rDNA intergenic region from *Arabidopsis thaliana*. Structural analysis, intraspecific variation and functional implications. *J Mol Biol* **221**: 1209-1222.

Grummt I. 2003. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev* **17**: 1691–1702.

Grummt I. 2007. Different epigenetic layers engage in complex crosstalk to define the epigenetic state of mammalian rRNA genes. *Hum Mol Genet* **15**:R21-27.

Hancock R. 2000. A new look at the nuclear matrix. *Chromosoma* **109**: 219–225.

Hanson R.E., Islam-Faridi M.N., Percival E.A., Crane C.F., Ji Y., McKnight T.D. 1996. Distribution of 5S and 18S-28S rDNA loci in a tetraploid cotton (Gossypium hirsutum L.) and its putative diploid ancestors. *Chromosoma* **105**: 55–61.

Hasterok R., Wolny E., Hosiawa M., Kowalczyk M., Kulak-Ksiazczyk S., Ksiazczyk T. *et al*. 2006. Comparative Analysis of rDNA Distribution in Chromosomes of various Species of Brassicaceae. *Ann Bot* **97**: 205–216.

Havecker E.R., Gao X., Daniel F Voytas D.F. 2004. The diversity of LTR retrotransposons. *Genome Biol* **5**: 225-225.6.

He P., Friebe B., Gill B., Zhou J-M. 2003. Allopolyploidy alters gene expression in the highly stable hexaploid wheat. *Plant Mol Biol* **52**: 401-414.

Heitz E. 1928. Das Heterochromatin der Moose I. *Jahrbuecher Wiss Bot* **69**: 762–818.

Heliot L., Kaplan H., Lucas L., Klein C., Beorchia A., *et al*. 1997. Electron tomography of metaphase nucleolar organizer regions: evidence for a twisted-loop organization. *Mol Biol Cell* **8**: 2199–2216.

Heng H.Q., Goetze S., Ye C.J., Liu G., Stevens J.B., Bremer S.W., Wykes S.M., Bode J., Krawetz S.A. 2004. Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *J Cell Sci* **117**: 999-1008.

Henikoff S., Ahmad K. 2005. Assembly of variant histones into chromatin. *Annu Rev Cell Dev Biol* **21**:133–153.

Henze K., Martin W. 2001. How do mitochondrial genes get into the nucleus? *Trends Genet* **17**: 383-387.

Heslop-Harrison J.S. 2000. Comparative Genome Organization in Plants: From Sequence and Markers to Chromatin and Chromosomes. *Plant Cell* **12**: 617–635.

Heslop-Harrison J.S. 2003. Planning for remodelling: nuclear architecture, chromatin and chromosomes. *Trends Plant Sci* **8**: 195-197.

Highett M.I., Beven A.F., Shaw P.J. 1993. Localization of 5S genes and transcripts in *Pisum sativum* nuclei. *J Cell Sci* **104**: 1151-1158.

Hizume M., Shibata F., Matsusaki Y., Garajova Z. 2002. Chromosome identification and comparative karyotypic analyses of four *Pinus* species. *Theor Appl Genet* **105**: 491–497.

Horn P.J., Peterson C.L. 2006. Heterochromatin assembly: A new twist on an old model. *Chromosome Res* **14**:83-94.

Hou D. 1971. Chromosome numbers of Trigonobalanus verticllata Forman (Fagaceae). *Acta Bot Neerlandica* **20**:543–549.

Houben A., Demidov D., Gernand D., Meister A., Leach C.R., Schubert I. 2003. Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. *Plant J* **33**: 967-973.

Huang S., Rothblum L.I., Chen D. 2006. Ribosomal chromatin organization. *Biochem Cell Biol* **84**: 444-449.

Hubbell H.R. 1985. Silver staining as an indicator of active ribosomal genes. *Stain Technol.* **60**: 284 -294.

Hueros G., Loarce Y., Ferrer E. 1993. A structural and evolutionary analysis of a dispersed repetitive sequence. *Plant Mol Biol* **22**: 635-643.

Ingle J., Timmis J., Sinclair J. 1975. The relationship between satellite DNA, ribosomal RNA redundancy and genome size in plants. *Plant Physiol* **55**: 496–501.

Jackson D.A. 2003. The anatomy of transcription sites. *Curr Opin Cell Biol* **15**: 311-317.

Jackson J.P., Lindroth A.M., Cao X., Jacobsen S.E. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**: 556–560.

Jacob S.T. 1995. Regulation of ribosomal gene transcription. *Biochem J* **306**: 617-626.

Janousek B., Zluvova J., Vyskot B. 2000 Histone H4 acetylation and DNA methylation dynamics during pollen development. *Protoplasma* **211**: 116 122.

Janousek B., Matsunaga S., Kejnovsky E., Zluvova J., Vyskot B. 2002. DNA methylation analysis of a male reproductive organ specific gene (MROS1) during pollen development. *Genome* **45**: 930-937.

Jasencakova Z., Meister A., Walter J., Turner B.M., Schubert I. 2000. Histone H4 acetylation of euchromatin and heterochromatin is cell cycle dependent and correlated with replication rather than with transcription. *Plant Cell* **12**:2087–2100.

Jasencakova Z., Meister A., Schubert I. 2001. Chromatin organization and its relation to replication and histone acetylation during the cell cycle in barley. *Chromosoma* **110**: 83–92.

Jasencakova Z., Soppe W.J., Turner B.M., Schubert I. 2003. Histone modifications in Arabidopsis– high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *Plant J* **33**: 471-480.

Jaynes R.A. 1961. Genetic and cytological studies in the genus *Castanea*.Ph.D. Dissertation. Yale Univ. pp 98

Jaynes R.A. 1962. Chesnut chromosomes. *Forest Sci* **8**: 372-377.

Jenuwein T. 2001. Re-SET-ing heterochromatin by histone methyltransferases. Trends Cell Biol **11**: 266-273.

Jenuwein T., Allis C.D. 2001. Translating the histone code. *Science* **293**:1074–1080.

Jiang J., Gill B.S. 2006. Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome* **49**: 1057-1068.

John H.A., Birnstiel M.L., Jones K.W. 1969. RNA-DNA hybrids at the cytological level. *Nature* **223**: 582-587.

Johnson G.P. 1988. Revision of *Castanea* Sect. Balanocastanon (Fagaceae). *J Arnold Arboretum* **69**: 25–49.

Johnsson H. 1946. Chromosome numbers of twin plants of *Quercus robur* and *Fagus sylvatica*. *Hereditas* 32: 469-472.

Jones J.H. 1986. Evolution of the Fagaceae: The implications of foliar features. *Ann Missouri Bot Garden* **73**: 228-275.

Jorgansen R.A., Cluster P.D. 1988. Modes and tempos in the evolution of nuclear ribosomal DNA: new characters for evolutionary studies and new markers for genetic and population studies. *Ann Mol Bot Gard* **75**: 1238-1247.

Kakutani T., Jeddeloh J.A., Richards E.J. 1995. Characterization of an *Arabidopsis thaliana* DNA hypomethylation mutant. *Nucl Acids Res* **23**: 130–137.

Kalendar R., Tanskanen J., Chang W., Antonius K., Sela H., Peleg O., Schulman A.H. 2008. Cassandra retrotransposons carry independently transcribed 5S RNA. *Proc Natl Acad Sci USA* **105**: 5833-5838.

Kamiya K., Harada K., Clyde M.M., Mohamed A.L. 2002. Genetic variation of *Trigonobalanus verticillata*, a primitive species of Fagaceae, in Malaysia revealed by chloroplast sequences and AFLP markers. *Genes Genet Syst* **77**:177-86.

Kang X.Y. 2002. Mechanism of 2n pollen occurring in Chinese white poplar. *J Beijing For Univ* **24**: 67–70.

Kankel M.W., Ramsey D.E., Stokes T.L., Flowers S.K., Haag J.R., Jeddeloh J.A. *et al*. 2003. Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* **163**:1109-1122.

Karvonen P., Karjalainen M., Savolainen O. 1993. Ribosomal RNA genes in scots pine (*Pinus sylvestris* L.): Chromosomal organization and structure. *Genetica* **88**: 59-68.

Kashkush K., Feldman M., Levy A.A. 2002. Gene loss, silencing, and activation in a newly synthesized wheat allotetraploid. *Genetics* **160**: 1651-1659.

Kiefer-Meyer M.C., Reddy A.S., Delseny M. 1996. Complex arrangement of dispersed repeated DNA sequences in *Oryza officinalis*. *Genome* **39**: 183-190.

Kim S.M., Dubey D.D., Huberman J.A. 2003. Early-replicating heterochromatin. *Genes Dev* **17**: 330-335.

Kim J.S, Klein P.E., Klein R.R., Price H.J., Mullet J.E., Stell D.M. 2005. Chromosome identification and nomenclature of *Sorghum bicolour*. *Genetics* **169**: 1169-1173.

King G.J. 2002. Through a genome, darkly: comparative analysis of plant chromosomal DNA. *Plant Mol Biol* **48**: 5-20.

Kobayashi T., Ganley A.R. 2005. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* **309**:1581-1584.

Kulikova O., Gualtieri G., Geurts R., Kim D.J., Cook D. *et al*. 2001. Integration of the FISH pachytene and genetic maps of *Medicago truncatula*. *Plant J* **27**: 49-58.

Kumar A., Pearce S.R., McLean K., Harrison G., Heslop- Harrison J.S., *et al*.1997. The Ty1-copia group of retrotransposons in plants: genomic organization, evolution, and use as molecular markers. *Genetica* **100**: 205–217.

Kuo M.H., Allis C.D. 1998. Roles of histone acetyltransferases and deacetylases in gene regulation. *BioEssays* **20**: 615–626.

Kuo H.F., Olsen K.M., Richards E.J. 2006. Natural Variation in a Subtelomeric Region of Arabidopsis: Implications for the Genomic Dynamics of a Chromosome End. *Genetics* **173**: 401–417.

Labrador M., Corces V.G. 2002 Setting the boundaries of chromatin domains and nuclear organization. *Cell* **111**: 151-154.

Lang P., Dane F., Kubisak T.L. 2006. Phylogeny.of *Castanea* (Fagaceae) based on chloroplast trnT-L-F sequence data. *Tree Genetics & Genomes* **2**: 132–139.

Lang P., Dane F., Kubisiak T., Huang H. 2007. Molecular evidence for an Asian origin and a unique westward migration of species in the genus *Castanea* via Europe to North America. *Mol Phylogent Evol* **43**: 49-59.

Lawrence R.J., Earley K., Pontes O., Silva M., Chen Z.J., Neves N. *et al.* 2004. A Concerted DNA Methylation/Histone Methylation Switch Regulates rRNA Gene Dosage Control and Nucleolar Dominance. *Mol Cell* **13**: 599-609.

Lee D.Y., Hayes J.J., Pruss D., Wolffe A.P. 1993. A positive role for histone acetylation in transcription factor access to nucleossomal DNA. *Cell* **72**: 73–84.

Lefort F., Lally M., Thompson D., Douglas G.C. 1998. Morphological traits, microsatellite fingerprinting and genetic relatedness of a stand of elite oaks (*Q. robur* L.) at Tullynally, Ireland. *Silvae Genetica* **47**: 257–262.

Lefort F., Douglas G.C. 1999. Occurrence and detection of triploids by microsatellite analysis. In: *Strategies for improvement of forest tree species.Proceedings of the Teagasc/TDC Symposium on Forest Genetics, COFORD* (Douglas G.C. ed). Dublin, Ireland, pp 19–35.

Lefort F., Douglas G.C., Thompson D. 2000. Microsatellite DNA profililng of phenotypically selected clones of Irish oak (*Quercus* spp.) and ash (Fraxinus exxelsior L.). *Silvae Genetica* **49**: 21–28.

Leister D. 2005. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends Genet* **21**: 655-663.

Levan A., Fredga K., Sandberg A.A. 1964. Nomenclature for centromeric position on chromosomes. *Hereditas* **52**: 201–220.

Levin D.A., Wilson A.C. 1976. Rates of evolution in seed plants: Net increase in diversity of chromosome numbers and species numbers through time. *Proc Natl Acad Sci USA* **73**: 2086–2090.

Levin D.A. 1983. Polyploidy and novelty in flowering plants. *Am Nat* **122**: 1–24.

Lewis J.D., Song Y., de Jong M.E., Bagha S.M., Ausio J. 2003. A walk though vertebrate and invertebrate protamines. *Chromosoma* **111**: 473– 482.

Li R.Q., Chen Z.D., Lu A.M., Soltis D.E., Soltis P.S., Manos P.S. 2004. Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *Int J Plant Sci* **165**: 311–324.

Lima-de-Faria A., Jaworska H. 1986. Late DNA synthesis in heterochromatin. *Nature* **217**: 138-142.

Lin X., Kaul S., Rounsley S., Shea T.P., Benito M.I., Town C.D. *et al*. 1999. Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**: 761-768.

Linhart Y.B. 1999. Variation in woody plants: molecular markers, evolutionary processes and conservation biology. In: *Molecular biology of woody plants. Forestry Sciences, Volume 64* (Jain S.M., Minocha S.C. eds) Kluwer Academic Publishers, The Netherlands, pp. 341-374.

Lippman Z., Martienssen R. 2004. The role of RNA interference in heterochromatic silencing. *Nature* **431**: 364-370.

Liu Z-L., Zhang D., Hong D-Y., Wang X-R. 2003. Chromosomal localization of 5S and 18S-5·8S-25S ribosomal DNA sites in five Asian *Pinus* species using fluorescence *in situ* hybridization. *Theor Appl Genet* **106**: 198–204.

Loyola A., Bonaldi T., Roche D., Imhof A., Almouzni G. 2006. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol Cell* **24**: 309–316.

Loudon J. 1838. Arboretum et fruticetum Botanicum. Longman, Rees, Orme, Brown and Green, London

Loureiro J., Rodriguez E., Dolezel J., Santos C. 2007. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37species. *Ann Bot* **100**: 875–888.

Lozano C.F., Hernandez C., Henao J.E. 1979. Hallazgo del genero *Trigonobalanus* Forman 1962 (Fagaceae) en el Neotropica. *Caldasia* **12**: 517–537.

Lukens L.N., Pires J.C., Leon E., Vogelzang R., Oslach L., Osborn T. 2006. Patterns of sequence loss and cytosine methylation within a population of newly resynthesized Brassica napus allopolyploids. *Plant Physiol* **140**: 336-348.

Luschnig C., Bachmair A., Schweizer D. 1993. Intraspecific length heterogeneity of the rDNA-IGR in *Arabidopsis thaliana* due to homologous recombination. *Plant Mol Biol* **22**: 543-545.

Ma J., Devos K.M., Bennetzen J.L. 2004. Analyses of LTR retrotransposon structures reveal recent and rapid genomic DNA loss in rice. *Genome Res* **14**: 860-869.

Malagnac F., Bartee L., Bender L. 2002. An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J* **21**: 6842–6852.

Manos P.S., Steele K.P. 1997. Phylogenetic analyses of ‘‘higher’’ Hamamelidiae based on plastid sequence data. *Am J Bot* **84**: 1407–1419.

Manos P.S., Doyle J.J., Nixon K.C. 1999. Phylogeny, Biogeography, and Processes of Molecular Differentiation in *Quercus* Subgenus *Quercus* (Fagaceae). *Mol Phylogenet Evol* **12**: 333–349.

Manos P.S., Zhou Z.K., Cannon C.H. 2001. Systematics of Fagaceae: phylogenetic tests of reproductive trait evolution. *Int J Plant Sci* **162**:1361-1379.

Marchler-Bauer A., Anderson J.B., Derbyshire M.K., DeWeese-Scott C., Gonzales N.R., Gwadz M. et al. 2007. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res* **35**: D237-240.

Martienssen R.A., Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science* **293**: 1070-1074.

Mayer K., Schüller C., Wambutt R., Murphy G., Volckaert G., Pohl T. *et al*. 1999. Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*. *Nature* **402**: 769: 777.

McClintock B. 1934. The relationship of a particular chromosomal element to the development of the nucleoli in Zea mays. *Z. Zellforch. Mikrosk. Anat.* **21**: 294 -328.

McClintock B. 1941. The stability of broken ends of chromosomes in *Zea mays*. *Genetics* **26**: 279–282.

McCormick S. 1993. Male Gametophyte Development. *Plant Cell* **5**: 1265-1275.

McStay B. 2006. Nucleolar dominance: a model for rRNA gene silencing. *Genes Dev* **20**: 1207–1214.

McStay B., Grummt I. 2008. The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu Rev Cell Dev Biol* **24**: 131–157.

Mendes M.C. 2007. O sobreiro ao longo dos tempos. In: *Árvores e florestas de Portugal. Os Montados. Muito para além das árvores*. (Silva J.S. ed) Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 77-106.

Meudt H.M., Clarke A.C. 2007. Almost Forgotten or Latest Practice? AFLP applications, analyses and advances. *Trends Plant Sci* **12**: 106-117.

Michalowski S.M., Allen G.C., Hall G.E. Jr., Thompson W.F., Spiker S. 1999. Characterization of randomly-obtained matrix attachment regions (MARs) from higher plants. *Biochemistry* **38**: 12795–12804.

Mok D.W.S., Peloquin S.J. 1975. Three mechanisms of 2n pollen formation in diploid potatoes. *Can J Genet Cytol* **17**: 217- 225.

Morais-Cecílio L., Delgado M., Jones R., Viegas W. 1996. Painting rye B chromosomes in wheat: Interphase chromatin organization, nucleolar disposition and association in plants with two, three or four Bs. *Chromosome Res* **4**: 195–200.

Morais-Cecílio L., Delgado M., Jones R.N., Viegas W. 2000. Modification of wheat rDNA loci by rye B chromosomes: a chromatin organization model. *Chromosome Res* **8**: 341–351.

Morales-Ruiz T., Ortega-Galisteo A.P., Ponferrada-Marin M.I., Ariza R.R., Roldan-Arjona T. 2006. DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA* **103**: 6853–6858.

Moss T., Stefanovsky V.Y. 1995. Promotion and regulation of ribosomal transcription in eukaryotes by RNA polymerase I. *Prog Nucleic Acid Res Mol Biol* **50**: 25-66.

Moss T., Stefanovsky V.Y. 2002. At the center of eukaryotic life. *Cell* **109**: 545-548.

Muir G., Fleming C.C., Schlotterer C. 2000 Species status of hybridising oaks. *Nature* **405**:1016.

Mukai Y., Endo T.R., Gill B.S. 1990. Physical mapping of the 5S rRNA multigene family in common wheat. *J Hered* **81**: 290–295.

Mukai Y., Endo T.R. Gill B.S. 1991. Physical mapping of the 18S.26S rRNA multigene family in common wheat: identification of a new locus. *Chromosoma* **100**: 71-78.

Müller A.,Marins M., Kamisugi Y., Meyer P. 2002. Analysis of hypermethylation in the RPS element suggests a signal function for short inverted repeats in *de novo* methylation. *Plant Mol Biol* **48:** 383–399.

Mullis K.B., Faloona F.A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* **155**: 335-350.

Nakayama S., Fukui K. 1997. Quantitative chromosome mapping of small plant chromosomes by improved imaging on CHIAS II. *Genes Genet Syst* **72**: 35-40.

Naujoks G., Hertel H., Ewald D. 1995. Characterization and propagation of an adult triploid pedunculate oak (*Quercus robur* L.). *Silvae Genetica* **44**: 282–286.

Navratilova A., Neumann P., Macas J. 2003. Karyotype analysis of four *Vicia* species using in situ hybridization with repetitive sequences. *Ann Bot* **91**: 921–926.

Neves N., Delgado M., Silva M., Caperta A., Morais-Cecílio L., Viegas W. 2005a. Ribosomal DNA heterochromatin in plants *Cytogenet Genome Res* **109**: 104–111.

Neves N., Viegas W., Pikaard C.S. 2005b. Nucleolar dominance and rRNA gene dosage control: a paradigm for transcriptiona regulation via an epigenetic on/off switch. In: *Plant Epigenetics Annual Plant Reviews, Vol. 19* (Meyer P. ed) Blackwell Publishing Ltd, pp 201-213.

Nickerson J. 2001. Experimental observations of a nuclear matrix. *J Cell Sci* **114**: 463-474.

Nixon K.C. 1989. Origins of Fagaceae. In: *Evolution systematics and fossil history of the Hamamelidae. Vol 2. "Higher" Hamamelidae* (Crane P.R., Blackmore S. eds) Clarendon, Oxford, pp 23-44

Nixon K.C., Crepet W.L. 1989. *Trigonobalanus* (Fagaceae):Taxonomic status and phylogenetic relationships. *Am J Bot* **76**: 828–841.

Nixon K.C. 1993. Infrageneric classification of *Quercus* (Fagaceae) and typification of sectional names. *Ann Sci For* **50**: 25-34.

Oakeley J., Podesta A., Jost J.P. 1997. Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation. *Proc Natl Acad Sci* **94**: 11721–11725.

Ohri D., Ahuja M.R. 1990. Giemsa C-banding karyotype in *Quercus* L. (Oak). *Silvae Genetica* **39**: 216–219.

Ohri D., Ahuja M.R. 1991. Giemsa C-banding in *Fagus sylvatica* L., *Betula pendula* Roth and *Populus tremula* L. *Silvae Genetica* **40**: 72–75.

Okada T., Singh M.B., Bhalla P.L. 2006. Histone H3 variants in male gametic cells of lily and H3 methylation in mature pollen. *Plant Mol Biol* **62**: 503–512.

Olszewska M.J., Osiecka R. 1984. The relationship between 2C DNA content, systematic position, and the level of nuclear DNA endoreplication during differentiation of root parenchyma in some dicotyledonous shrubs and trees - comparison with herbaceous species. *Biochemie Und Physiologie Der Pflanzen* **179**: 641-657.

Oono K., Sugiura M. 1980. Heterogeneity of the ribosomal RNA gene clusters in rice. *Chromosoma* **76**: 85–89.

Paffetti D., Vettori C., Caramelli D., Vernesi C., Lari M., Paganelli A. *et al*. 2007. Unexpected presence of Fagus orientalis complex in Italy as inferred from 45,000-year-old DNA pollen samples from Venice lagoon. *BMC Evol Biol* **7**(Suppl 2):S6.

Paiva J. 2007. O castanheiro, uma perspectiva histórica perspectivas. In: *Árvores e florestas de Portugal. Do Castanheiro ao Teixo. As outras espécies florestais*. (Silva J.S. ed). Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 39-50.

Pan G., Zhou Y., Fowke L.C., Wang H. 2004. An efficient method for flow cytometric analysis of pollen and detection of 2n nuclei in *Brassica napus* pollen. *Plant Cell Rep* **23**: 196–202.

Paszko B. 2006. A critical review and a new proposal of karyotype asymmetry indices. *Pl Syst Evol* **258**: 39–48.

Pearce S.R., Pich U., Harrison G., Flavell A.J., Heslop Harrison J.S. 1996. The Ty1-copia group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Res* **4**: 357–364.

Pedrosa-Harand A., de Almeida C.C., Mosiolek M., Blair M.W., Schweizer D., Guerra M. 2006. Extensive ribosomal DNA amplification during Andean common bean (*Phaseolus vulgaris* L.) evolution. *Theor Appl Genet* **112**: 924–933.

Pereira H. 2007. A cortiça - um material único. In: *Árvores e florestas de Portugal. Os Montados. Muito para além das árvores*. (Silva J.S. ed) Público, Comunicação Social, SA., Fundação Luso-Americana para o Desenvolvimento, pp 59-75.

Pérez-Prieto L.J., López-Roca J.M., Martinez-Cutillas A., Minguez P.F., Gómez-Plaza E. 2002. Maturing wines in oak barrels. Effects of origin, volume, and age of the barrel on the wine volatile composition. *J Agric Food Chem* **22**:3272-3276.

Petit R.J., Bodénès C., Ducousso A., Roussel G., Kremer A. 2003. Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**: 151–164.

Petit R., Hampe A. 2006. Some evolutionary consequences of being a tree. *Ann. Rev. Ecol. Evol. Syst*. **37**: 187-214.

Pikaard C. 2000. Nucleolar dominace: uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant Mol Biol* **43**: 163-177.

Pikaard C.S. 2002. Transcription and tyranny in the nucleolus: the organization, activation, dominance, and repression of ribosomal RNA genes. In: The Arabidopsis Book (Somerville C.R., Meyerowitz E.M. eds) American Society of Plant Biologists, Rockville MD.

Pikaard C.S., Lawrence R.J. 2002. Uniting the paths to gene silencing. *Nat Genet*. **32**: 340-341.

Pradhan S., Adams R.L.P. 1995. Distinct CG and CNG DNA methyltransferases in *Pisum sativum*. *Plant J* **7**: 471-481.

Prescott D.M. 1964. Cellular sites of RNA synthesis. *Prog Nucl Acids Res* **3**: 33-57.

Preuss S., Pikaard C.S. 2007. rRNA gene silencing and nucleolar dominance: insights into a chromosome-scale epigenetic on/off switch. *Biochim Bioph Acta* **1769**: 383–392.

Prokopowich C.D., Gregory T.R., Crease T.J. 2003. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **46**: 48-50.

Rabinowicz P.D., Bennetzen J.L. 2006. The maize genome as a model for efficient sequence analysis of large plant genomes. *Curr Opin Plant Biol* **9**: 149-156.

Rabl C. 1885. Über zelltheilung. *Morphol Jarbuch* **10**: 214-330.

Raskina O., Belyayev A., Nevo E. 2004a. Activity of the En/Spm like transposons in meiosis as a base for chromosome repatterning in a small, isolated, peripheral population of *Aegilops speltoides* Tausch. *Chromosome Res* **12**: 153–161.

Raskina O., Belyayev A., Nevo E. 2004b. Quantum speciation in *Aegilops*: molecular cytogenetic evidence from rDNA cluster variability in natural populations. *Proc Natl Acad Sci USA* **101**: 14818–14823.

Raven P.H., Evert R.F., Eichhorn S.E. 1999. Meiosis and sexual reproduction. In: *Biology of Plants* (Cloud D. ed) New York: WH Freeman and company, Worth publishers pp 169-182.

Reamon-Buttner S.M., Schmidt T., Jung C. 1999. AFLPs represent highly repetitive sequences in *Asparagus officinalis* L. *Chromosome Res* **7**: 297-304.

Reddy P., Appels R. 1989. A second locus for the 5S multigene family in *Secale* L.: sequence divergence in two lineages of the family. *Genome* **32**: 456–467.

Redi C.A., Garagna S., Zacharias H., Zuccotti M., Capanna E. 2001. The other chromatin. *Chromosoma* **110**: 136–147.

Reuter G., Fisher A., Hofman I. 2005. Heterochromatin and the control of gene silencing in plants. In: *Plant Epigenetics. Annual Plant Reviews*, Volume 19. (Meyer P ed). Blackwell Publishing, pp. 106-133.

Ribeiro T., Barão A., Viegas W., Morais-Cecílio L. 2008a. Molecular cytogenetics of forests trees. *Cytogenet Genome Res* **120**: 220-227.

Ricchetti M., Fairhead C., Dujon B. 1999. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* **402**: 96–100.

Richards E.J., Elgin S.C. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* **108**: 489–500.

Richly E., Leister D. 2004a. NUMTs in sequenced eukaryotic genomes. *Mol Biol Evol* **21**: 1081–1084.

Richly E., Leister D. 2004b. NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. *Mol Biol Evol* **21**:1972–1980.

Riddle N.C., Richards E.J. 2002. The control of natural variation in cytosine methylation in Arabidopsis. *Genetics* **162**: 355–363.

Rogers S.O., Bendich A.J. 1987. Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol Biol* **9**: 509-520.

Ronemus M.J., Galbiati M., Ticknor C., Chen J., Dellaporta S.L. 1996. Demethylation-induced developmental pleiotropy in Arabidopsis. *Science* **273**: 654–657.

Roussel P., Andre C., Comai L., Hernandez-Verdun D. 1996. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *J Cell Biol* **133**: 235–246.

Rushton B.S. 1993. Natural hybridization within the genus *Quercus* L. *Ann Sci For* **50**: 73–90.

Salomon S., Puchta H. 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *EMBO J* **17**: 6086–6095.

SanMiguel P., Tikhonov A., Jin Y.-K., Motchoulskaia N., Zakharov D., Melake-Berhan A., Springer P.S., Edwards K.J., Lee M., Avramova Z., Bennetzen J.L. 1996. Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**: 765–768.

SanMiguel P., Bennetzen J.L. 1998. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. *Ann Bot* **82**: 37-44.

SanMiguel P., Gaut B.S., Tikhonov A., Nakajima Y., Bennetzen J.L. 1998. The paleontology of intergene retrotransposons of maize. *Nat Genet* **20**: 43-45.

Santoro R., Li J., Grummt I. 2002. The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nature Genet* **32**: 393-396.

Santoro R. 2005. The silence of the ribosomal RNA genes. *Cell Mol Life Sci* **62**: 2067-2079.

Santoro R., Grummt I. 2005. Epigenetic mechanism of rRNA gene silencing: Temporal order of NoRCmediated histone modification, chromatin remodeling, and DNA methylation. *Mol Cell Biol* **25**: 2539-2546.

Saze H., Scheid O.M., Paszkowski J. 2003. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* **34**: 65–69.

Scalfi M., Troggio M., Piovani P., Leonardi S., Magnaschi G., Vendramin G., Menozzi P. 2004. A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). *Theor Appl Genet* **108**: 433–441.

Schaal B.A., Learn G.H.J. 1988. Ribosomal DNA variation with and among plant populations. *Ann Mo Bot Gard* **75**: 1207-1216.

Scheid O.M., Jakovleva L., Afsar K., Maluszynska J., Paszkowski J. 1996. A change in ploidy can modify epigenetic silencing. *Proc Natl Acad Sci USA* **93**: 7114–7119.

Schmidt T., Heslop-Harrison J.S. 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci* **3**: 195-199.

Schroeder-Reiter E., Houben A., Grau J., Wanner G. 2006. Characterization of a peg-like terminal NOR structure with light microscopy and high-resolution scanning electron microscopy. *Chromosoma* **115**: 50–59.

Schwarz O. 1964. *Quercus* L.. In: *Flora Europaea 1, 1st ed.* (Tutin T.G., Heywood V.H., Burges N.A., Morre D.M., Valentine D.H., Walters S.M., Webb D.A. eds.) Cambridge University Press, Cambridge, pp. 72–76.

Schwarzacher T., Heslop-Harrison J.S. 2000. *Practical in situ hybridization*. Bios, Oxford.

Sharma S., Raina S.N. 2005. Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. *Cytogenet Genome Res* **109**: 15-26.

Shaw P.J., Jordan E.G. 1995. The nucleolus. *Annu Rev Cell Dev Biol* **11**: 93–121.

Sheldon C.C., Conn A.B., Dennis E.S., Peacock W.J. 2002. Different regulatory regions are required for the vernalization-induced repression of FLOWERING LOCUS C and for the epigenetic maintenance of repression. *Plant Cell* **14**: 2527-2537.

Shen C. F. 1992. A monograph of the genus *Fagus* Tourn. ex L. (Fagaceae). Ph. D. dissertation, The City University of New York, New York, USA.

Sherald J.L., Santamour F.S., Hajela R.K. Jr., Hajela N., Sticklen M.B. 1994. A Dutch elm disease resistant triploid elm. *Can J Forest Res* **24**: 647–653.

Shubert I. 2007. Chromosome evolution. *Curr Opin Plant Biol* **10**: 109–115.

Siljak-Yakovlev S., Cerbah M., Coulaud J., Stoian V., Brown S. *et al*. 2002. Nuclear DNA content, base composition, heterochromatin and rDNA in *Picea omorika* and *Picea abies*. *Theor Appl Genet* **104**: 505-512.

Sims R.J., Nishioka K., Reinberg D. 2003. Histone lysine methylation: a signature for chromatin function. *Trends in Genet* **19**: 629-639.

Slocum M.K., Figdore S.S., Kennard W.C., Suzuki J.Y., Osborn T.C. 1990. Linkage arrangement of restriction fragment length polymorphism loci in *Brassica oleracea*. *Theor Appl Genet* **80**: 57–64.

Smyth D.R. 1991. Dispersed repeats in plant genomes. *Chromosoma* **100**: 355–359.

Snowdon R.J., Friedrich T., Friedt W., Kohler W. 2002. Identifying the chromosomes of the A and C genome diploid Brassica species *B. rapa* and *B. oleracea* in their amphidiploid *B. napus*. *Theor Appl Genet* **104**: 533-538.

Soepadmo E. 1968. A revision of the genus *Quercus* L. subgen. *Cyclobalanopsis* (Oersted) Schneider. *Malesia Gard Bull Sing* **22**: 355–427.

Soepadamo E. 1972. Fagaceae In: *Flora Malesiana*, series I, volume 7 (Van Steenis C.G. ed) Noordhoff International Publishing, Leyden, pp. 265–388.

SoppeW.J., Jasencakova Z., Houben A., Kakutani T., Meister A., Huang M.S. et al. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *EMBO J* **21**: 6549–6559.

Stack S.M. 1991. Staining plant cells with silver. II. Chromosome cores. *Genome* **34**: 900-908.

Stebbins G.L. 1971. Chromosome Evolution. In: Higher Plants. Addison-Wesley Publishing Company, Massachusetts.

Sterck L., Rombauts S., Jansson S., Sterky F., Rouzé P., van de Peer Y. 2005. EST data suggest that poplar is an ancient polyploid. *New Phytol* **167**: 165–170.

Strohner R., Nemeth A., Jansa P., Hofmann-Rohrer U., Santoro R., Langst G., Grummt I. 2001. NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines. *Embo J* **20**: 4892–4900.

Stupar R.M., Lilly J.W., Town C.D., Cheng Z., Kaul S., Buell C.R., Jiang J. 2001. Complex mtDNA constitutes an approximate 620-kb insertion on Arabidopsis thaliana chromosome 2: implication of potential sequencing errors caused by large-unit repeats. *Proc Natl Acad Sci USA* **98**: 5099–5103.

Sun J.M., Chen H.Y., Espino P.S., Davie J.R. 2007. Phosphorylated serine 28 of histone H3 is associated with destabilized nucleosomes in transcribed chromatin. *Nucleic Acids Res* **35**: 6640–6647.

Talbert P.B., Masuelli R., Tyagi A.P., Comai L., Henikoff S. 2002. Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. *Plant Cell* **14**:1053-1066.

Tanabe H., Muller S., Neusser M., von Hase J., Calcagno E., Cremer M., Solovei I., Cremer C., Cremer T. 2002. Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc Natl Acad Sci USA* **99**: 4424-4429.

Tanaka I. 1997. Differentiation of generative and vegetative cells in angiosperm pollen. *Sex Plant Reprod* **10**: 1–7.

Tanaka I., Ono K., Fukuda T. 1998. The developmental fate of angiosperm pollen is associated with a preferential decrease in the level of histone H1 in the vegetative nucleus. *Planta* **206**: 561-569.

Tariq M., Saze H., Probst A.V., Lichota J., Habu Y., Paszkowski J. 2003. Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. *Proc Natl Acad Sci USA* **100**: 8823–8827.

Tariq M., Paszkowski J. 2004. DNA and histone methylation in plants. *Trends Genet* **20**: 244-251.

Tetko I.V., Haberer G., Rudd S., Meyers B., Mewes H.W., et al. 2006. Spatiotemporal expression control correlates with intragenic scaffold matrix attachment regions (S/MARs) in Arabidopsis thaliana. *PLoS Comput Biol* **2**(3): e21.

Tian H.Q., Russell S.D. 1998. The fusion of sperm cells and the function of male germ unit (MGU) of tobacco (*Nicotiana tabacum* L.). *Sex Plant Reprod* **11**: 171–176.

Tian L., Fong M.P., Wang J.J., Wei N.E., Jiang H., Doerge R.W., Chen Z.J. 2005. Reversible histone acetylation and deacetylation mediate genome-wide, promoter-dependent and locus-specific changes in gene expression during plant development. *Genetics* **169**: 337–345.

Tikhonov A.P., Bennetzen J.L., Avramova Z.V. 2000. Structural Domains and Matrix Attachment Regions along Colinear Chromosomal Segments of Maize and Sorghum. *Plant Cell* **12**: 249–264.

Tompa R., McCallum C.M., Deltrow J., Henikoff J.G., van Steensel B., Henikoff S. 2002. Genome-wide profiling of DNA methylation reveals transposon targets of CHROMOMETHYLASE3. *Curr Biol* **12**: 65-68.

Trelease W.1924. The American oaks. *Mem Natl Acad Sci* **20**: 1–255.

Tremethick D.J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* **128**: 651-654.

Tsai Y-T., Lin C-I., Chen H-K., Lee K-M., Hsu C-Y., Yang S-J., Yeh N-H. 2008. Chromatin tethering effects of hNopp140 are involved in the spatial organization of nucleolus and the rRNA gene transcription. *J Biomed Sci* **15**: 471–486.

Turner B.M. 2001. *Chromatin and gene regulation*. Blackwell Science Ltd., Oxford. Chapter 9 pp.175.

Yoder J.A., Walsh C.P., Bestor T.H. 1997. Cytosine methylation and the ecology of intragenomic plasmids. *Trends Genet* **13**: 335-340.

Yu X., Gabriel A. 1999. Patching broken chromosomes with extranuclear cellular DNA. *Mol Cell* **4**: 873–881.

Uchida W., Matsunaga S., Sugiyama R., Kawano S. 2002. Interstitial telomere-like repeats in the *Arabidopsis thaliana* genome. *Genes Genet Syst* **77**: 63-67.

Ueda K., Kinoshita Y., Xu Z., Ide N., Ono M., Akahori Y. *et al*. 2000 Unusual core histones specifically expressed in male gametic cells of *Lilium longiflorum*. *Chromosoma* **108**: 491–500.

Valbuena-Carabana M., Gonzalez-Martinez S.C., Sork V.L., Collada C., Soto A., Goicoechea P.G., Gil L. 2005. Gene flow and hybridisation in a mixed oak forest (*Quercus pyrenaica* Willd. and *Quercus petraea* (Matts.) Liebl.) in central Spain. *Heredity* **95**: 457-465.

Valdiviesso T., da Costa R.L. 2006. Polinizações Controladas em *Castanea* spp. e Caracterização da Descendência e Progenitores por Microsatélites. *Silva Lusitana* **14**:23 – 31.

van Driel R., Fransz P.F., Verschure P.J. 2003. The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci* **116**: 4067–4075.

van Driel R., Fransz P.F. 2004. Nuclear architecture and genome functioning in plants and animals: what can we learn from both? *Exp Cell Res* **296**: 86–90.

Van Valen L. 1976. Ecological species, multispecies and oaks. *Taxon* **25**: 233–239.

Vassetzky Y., Hair A., Mechali M. 2000. Rearrangement of chromatin domains during development in *Xenopus*. *Genes Dev* **14**: 1541-1552.

Vettese-Dadey M., Grant P.A., Hebbes T.R., Crane-Robinson C., Allis C.D., Workman J.L. 1996. Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA *in vitro*. *EMBO J* **15**: 2508– 2518.

Vershinin, A.V., Schwarzacher, T., Heslop-Harrison, J.S. 1995. The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. *Plant Cell* **7**: 1823– 1833.

Viegas W., Neves N., Silva M., Caperta A., Morais-Cecílio L. 2002. Nucleolar dominance: a "David and Goliath" chromatin imprinting process. *Curr Genomics* **3**: 563–576.

Vieira R., Queiroz A., Morais L., Barão A., Mello-Sampayo T., Viegas W. 1990. 1R chromosome nucleolus organizer region activation by 5-azacytidine in wheat-rye hybrids. *Genome* **33**: 707-712.

Vos P., Hogers R., Bleeker M., Reijans M., van de Lee T., Hornes M. et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.

Wang B.D., Butylin P., Strunnikov A. 2006a. Condensin Function in Mitotic Nucleolar Segregation is Regulated by rDNA Transcription. *Cell Cycle* **5**: 2260-2267.

Wang J., Tian L., Lee H.S., Wei N.E., Jiang H. *et al*. 2006b. Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* **172**: 507–17.

Wanner G., Formanek H. 1995. Imaging of DNA in human and plant chromosomes by high-resolution scanning electron microscopy. *Chromosome Res* **3**: 368–374.

Wendel J.F., Schnabel A., Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc Natl Acad Sci USA* **92**: 280–284.

Whittemore A.T., Schall B.A. 1991 Interspecific gene flow in sympatric oaks. *Proc Natl Acad Sci* **88**: 2540-2544.

Wicker T., Guyot R., Yahiaoui N., Keller B. 2003. CACTA Transposons in Triticeae. A Diverse Family of High-Copy Repetitive Elements. *Plant Physiol* **132**: 52–63.

Wouters-Tyrou D., Martinage A., Chevaillier P., Sautiere P. 1998. Nuclear basic proteins in spermiogenesis. *Biochimie* **80**: 117– 128.

Zhai J., Liu J., Liu B., Li P., Meyers B.C. *et al*. 2008. Small RNA-Directed Epigenetic Natural Variation in *Arabidopsis thaliana*. *PLoS Genet* **4**: e1000056. doi:10.1371/journal.pgen.1000056.

Zhang Y., Reinberg D. 2001. Transcription regulation by histonemethylation: interplay between different covalent modifications of the core histone tails. *Genes Dev* **15**: 2343-2360.

Zhang X., Wessler S.R. 2004. Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. *Proc Natl Acad Sci USA* **101**: 5589-5594.

Zhang Z., Kang X., Zhang P., Li Y., Wang J. 2007. Incidence and molecular markers of 2n pollen in *Populus tomentosa* Carr. *Euphytica* **154**: 145–152.

Zimmer E., Jupe E., Walbot V. 1988. Ribosomal gene structure, variation, and inheritance in maize and its ancestors. *Genetics* **120**: 1125-1136.

Zoldoš V., Papeš D., Brown S.C., Panaud O., Šiljak-Yakovlev S. 1998. Genome size and base composition of seven *Quercus* species: inter- and intra-population variation. *Genome* **41**: 162-168.

Zoldoš V., Papeš D., Cerbah M., Panaud O., Besendorfer V., Šiljakyakovlev S. 1999. Molecular-cytogenetic studies of ribosomal genes and heterochromatin reveal conserved genome organization among 11 Quercus species. *Theor Appl Genet* **99**: 969–977.

Zurita F., Jimenez R., Burgos M., Diaz de la Guardia R. 1998. Sequential silver staining and in situ hybridization reveal a direct association between rDNA levels and the expression of homologous nucleolar organizer regions: a hypothesis for NOR structure and function. *J. Cell Sci* **111**: 1433 - 1439.

# *Table of Contents*