UNIVERSIDADE TÉCNICA DE LISBOA

RNA interference and nucleolar dominance establishment in *Arabidopsis suecica*

Pedro Miguel Melo da Costa Nunes

ORIENTADOR: Doutora Maria Wanda Sarujine Viegas

JÚRI: *Presidente:*
Reitor da Universidade Técnica de Lisboa

Vogais:

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

- Doutor Miguel Godinho Ferreira, investigador principal do Instituto Gulbenkian de Ciência da Fundação Calouste Gulbenkian;

- Doutor Joao António Augusto Ferreira, professora auxiliar da Faculdade de Medicina da Universidade de Lisboa;

- Doutora Maria Margarida Cabrita Xavier Delgado, professora auxiliar da Universidade Lusófona de Humanidades e Tecnologias, na qualidade de especialista;

- Doutora Maria Leonor Mota Morais Cecílio, professor auxiliar do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

DOUTORAMENTO EM BIOLOGIA

LISBOA

2007

RNA interference and nucleolar dominance establishment in *Arabidopsis suecica*

Pedro Miguel Melo da Costa Nunes

ORIENTADOR: Doutora Maria Wanda Sarujine Viegas

JÚRI: *Presidente:*
    Reitor da Universidade Técnica de Lisboa

    Vogais:

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

    - Doutor Miguel Godinho Ferreira, investigador principal do Instituto Gulbenkian de Ciência da Fundação Calouste Gulbenkian;

    - Doutor Joao António Augusto Ferreira, professora auxiliar da Faculdade de Medicina da Universidade de Lisboa;

    - Doutora Maria Margarida Cabrita Xavier Delgado, professora auxiliar da Universidade Lusófona de Humanidades e Tecnologias, na qualidade de especialista;

    - Doutora Maria Leonor Mota Morais Cecílio, professor auxiliar do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

"Tese apresentada neste Instituto para obtenção do grau de doutor"

DOUTORAMENTO EM BIOLOGIA

LISBOA

2007

Resumo

Os smRNAs constituem uma peça fundamental dos mecanismos de regulação de expressão génica. Estes RNAs de 21-24nt (siRNAs) medeiam o estabelecimento de marcas epigenéticas repressivas da transcrição, dirigindo complexos efectores a sequências de DNA homólogas.

Neste trabalho pretende-se avaliar o papel desta via de modulação da cromatina no estabelecimento da dominância nucleolar. Este fenómeno de silenciamento génico, mediado por mecanismos epigenéticos, define a inactivação transcripcional preferencial da totalidade dos *loci* de DNA ribossomal provenientes de um dos progenitores em espécies poliplóides. O silenciamento dos genes de rDNA está correlacionado com o estabelecimento de marcas heterocromáticas nas suas regiões promotoras, nomeadamente metilação de DNA e modificações pós-traducão de histonas. Foram identificados siRNAs (24nts) homólogos ao promotor dos genes ribossomais, cuja biogénese depende da polimerase PolIV e de outras proteínas integrantes da via de siRNAs heterocromáticos em *Arabidopsis sp.*. A identificação de transcritos de rDNA provenientes do genoma de *A. thaliana*, normalmente silenciados no alopoliplóide *A. suecica* (*A. thaliana x A. arenosa*), em linhas "knock down" destes genes demonstra pela primeira vez o envolvimento de siRNAs no estabelecimento da dominância nucleolar.

Palavras-chave: rDNA, siRNAs, polIV, dominância nucleolar, *Arabidopsis sp.*

Abstract

The discovery of small RNAs has added a new level to our understanding of gene regulation mechanisms. Small interfering RNAs (siRNAs) are involved in the establishment of repressive epigenetic marks in a variety of organisms by guiding RNA induced silencing complexes to homologous DNA sequences.

Could the endogenous siRNA heterochromatic pathway be linked to the establishment of nucleolar dominance? The latter is a large scale manifestation of directed epigenetic gene silencing, resulting in the transcriptional inactivation of the entire rRNA gene cluster from one parent in a hybrid species. Inactivation of rRNA genes is linked to the establishment of heterochromatic marks, both at the DNA and histone levels, in the promoter region. We were able to identify 24nt siRNAs homologous to the rRNA gene promoter in Arabidopsis sp., whose biogenesis is dependent on nuclear RNA polymerase IV and other known proteins of the endogenous siRNA pathway in *Arabidopsis*. RNAi mediated knockdown of genes in this pathway in the natural allopolyploid hybrid *A. suecica* (*A. thaliana x A. arenosa*) disrupted nucleolar dominance, causing the production of rRNA transcripts from the underdominant *A. thaliana* rDNA genes. This observation implicates nuclear siRNAs in the establishment of nucleolar dominance in *A. suecica*.

Key words: rDNA, siRNAs, RNA Polymerase IV (pol IV), nucleolar dominance, *Arabidopsis sp.*

Resumo (português)

O fenómeno de dominância nucleolar é um dos mais importantes exemplos de modulação de expressão génica mediada por mecanismos epigenéticos. Ocorre em espécies poliplóides resultando no silenciamento preferencial dos *loci* de rDNA de uma das espécies parentais, sendo os genes de rDNA do complemento genómico do progenitor dominante responsáveis pela transcrição da totalidade do rRNA ribossomal no núcleo híbrido. Apesar de o fenómeno de dominância nucleolar ter sido identificado em várias espécies hibridas, tanto em plantas como animais, é ainda desconhecido o completo mecanismo responsável pelo estabelecimento da origem parental de rDNA silenciado.

Os genes de rDNA estão organizados em *loci* constituídos por centenas a milhares de cópias organizadas em série, sendo todas competentes para a transcrição do percursor dos RNAs ribossomais. Determinante na actividade transcripcional dos genes ribossomais, dentro de um mesmo *locus*, são as marcas epigenéticas associadas ao seu promotor (revisão em Pikaard, 1999). Em *Arabidopsis thaliana*, a cromatina dos promotores de genes activos é caracterizada por marcas características de eucromatina, nomeadamente baixos níveis de metilação de DNA e nucleossomas compostos por histonas acetiladas e histonas H3 tri-metilada na lisina 4 (H3^{3m}K4). Por sua vez, genes de rRNA inactivos apresentam elevados níveis de metilação do DNA, baixo nível de acetilação de histonas e a histona H3 di-metilada na lisina 9 (H3^{2m}K9), marcas epigenéticas características de heterocromatina (Lawrence *et al.*, 2003; Earley et *al.*, 2006).

No alopoliplóide natural *A. suecica* (**n**=26, *A. thaliana x A. arenosa*), a totalidade dos genes de rDNA do complemento genómico de *A. thaliana* encontram-se silenciados e associados a marcas heterocromáticas. Por sua vez, os *loci* de rDNA de *A. arenosa* apresentam uma fracção dos seus genes organizados como heterocromatina e a restante em eucromatina, sendo estes últimos os genes responsáveis pela transcrição do RNA ribossomal. O recurso a tratamentos com agentes que induzem a hipometilação do DNA (5-azacitidina) ou a hiperacetilação de histonas (tricostatina-A, inibidor químico de desacetilases de histonas) resulta na reactivação de genes ribossomais de *A. thaliana* no alopoliplóide, o que demonsta a importância destas marcas epigenéticas no estabelecimento da dominância nucleolar (Chen and Pikaard, 1997). A tecnologia de RNA de interferência permite reduzir substancialmente os níveis de produção de uma dada proteína alvo, através da degradação dirigida do seu RNA codificante, o que no contexto de um alopoliplóide possibilita tornear problemas originados por redundância génica (Lawrence *et al.,* 2003*)*. Recorrendo a esta tecnologia procedeu-se à identificação de duas desacetilases (HDT1 e HDA6) (Lawrence *et al.*, 2004; Earley *et al.*, 2006) e uma metiltranferase de DNA (DRM2) (Preuss *et al.*, em preparação) necessárias para o

estabelecimento da dominância nucleolar. A desacetilase HDA6 (Aufsatz *et al.*, 2002) e a metiltransferase DRM2 (Cao *et al.*, 2003) estão ambas implicadas no processo de silenciamento génico mediado por RNAi, um mecanismo endógeno baseado no estabelecimento localizado de marcas heterocromáticas dirigido por moléculas de RNA homólogas à sequência de DNA alvo.

A importância, no contexto da dominância nucleolar, de modificações epigenéticas e do papel das enzimas HDA6 e DRM2 levanta a hipótese dos mecanismos endógenos de RNA de interferência, mediados por RNAs não-codificantes de 21 a 24 nucleótidos (smRNAs), contribuirem para a determinação de qual o complemento de rDNA parental silenciado em espécies poliplóides.

Identificação e mapeamento de smRNAs homólogos a genes de rDNA em *Arabidopsis thaliana*.

Por forma a avaliar o envolvimento da via de smRNAs endógenos no estabelecimento da dominância nucleolar em *A. suecica*, procedeu-se à identificação, mapeamento e caracterização de smRNAs na espécie diplóide *A. thaliana*. Recorrendo às bases de dados de smRNAs disponíveis em *Arabidopsis sp.* (ASRP database www.asrp.cgrb.oregonstate.edu/db/; Gustafson *et al.*, 2005) foi possível identificar um elevado número destes RNAs não codificantes (21-24nt) homólogos à sequência génica de rDNA 45S. Curiosamente, observou-se uma distribuição discreta de smRNAs na região correspondente ao promotor dos genes ribossomais, tendo a validação por hibridação RNA:RNA (Northern blot) demonstrado ser a classe de 24nt a mais abundante. A biogénese destes smRNAs está dependente da acção da RNA polimerase RDR2 e da RNAse DCL3, duas proteínas associadas à via de siRNAs heterocromáticos (RNAs de interferência da classe de 24nt) (Xie et al., 2004) implicada na modulação da cromatina. Alelos nulos da RNA polimerase RDR6, associada à biogénese de smRNAs que actuam em *trans* ou na via de silenciamento de transgenes, não afectam os níveis de acumulação de siRNAs de 24nt. Finalmente, as RNAses DCL4, responsável pela biogénese de siRNAs (classe 21nt), DCL1, associada à biogénese de micro RNAs (21-24nt) e DCL2, envolvida na resposta a vírus ou a stresses abióticos, não estão envolvidas na acumulação dos siRNAs homólogos à zona do promotor dos genes ribossomais.

A produção de siRNAs endógenos implica a presença de um percursor de RNA que funciona como molde para a acção da RDR2 e posterior processamento do RNA de dupla cadeia daí resultante pela DCL3 em siRNAs de 24nts. Tendo sido identificados siRNAs tanto a montante como a jusante do promotor dos genes de rDNA, procedeu-se à amplificação de RNAs potencialmente envolvidos na sua biogénese por RT-PCR, recorrendo a uma combinação de primers que flanqueiam a região de acumulação de siRNAs. A sequenciação dos fragmentos de PCR obtidos demonstrou serem homólogos à zona do promotor, indicando que a unidade génica de rDNA é em si mesma responsável

pela produção dos siRNAs identificados e que estes não são o resultado da expressão de um outro *locus*.

Biogénese de siRNAs e seus efeitos na organização da cromatina nos loci de rDNA.

O projecto de sequenciação do genoma de *Arabidopsis sp.* revelou dois novos genes que, pela similaridade da sua sequência de DNA, poderiam codificar subunidades semelhantes às integrantes das RNA polimerases I, II e III, responsáveis pela transcrição da totalidade dos genes de um genoma. A sua função demonstrou que estas subunidades são constituintes de uma nova RNA polimerase específica de plantas cuja função é determinante na biogénese de siRNAs heterocromáticos (24nt), eliminados em plantas contendo alelos não funcionais destas subunidades. A ausência de siRNAs traduz-se na hipometilação de sequências de DNA alvo, como sejam os genes que codificam o RNA ribossomal 5S ou os retrotransposões AtSNI (Herr *et al.*, 2005; Onodera *et al.*, 2005). As duas novas subunidades, NRPD2a e NRPD1a, são parte integrante da agora denominada PolIVa. Esta nova RNA polimerase, bem como uma outra forma posteriormente identificada, PolIVb (NRPD2a+NRPD1b) (Kanno *et al.*, 2005; Pontier *et al.*, 2005), são membros da via envolvida na biogénese dos siRNA heterocromáticos, conjuntamente com a proteína remodeladora de cromatina DRD1 (Kanno *et al.*, 2004), a metiltransferase DRM2 (Cao *et al.*, 2003), a AGO4 (Zilberman *et al.*, 2003), proteína que se liga a siRNAs de 24nts e que dirige o "targeting" do complexo efector para a sequência de DNA alvo, e finalmente a RNA metilase HEN1 (Yang *et al.*, 2006), responsável pela estabilização dos smRNAs. Todas estas proteínas demonstraram estar envolvidas na biogénese dos siRNAs homólogos à região promotora do rDNA, colocando-os como potenciais factores envolvidos no estabelecimento da cromatina ribossomal. Curiosamente, um outro smRNA homólogo à região intergénica dos genes de rDNA, siR759, apresenta um mecanismo de biogénese diverso do acima descrito. A acção das proteínas DRD1 e DRM2, neste caso, é substituída pela proteína remodeladora de cromatina DDM1 e pela metiltransferase MET1. Estas últimas estão envolvidas no estabelecimento de marcas heterocromáticas transmissíveis, como a metilação de manutenção do DNA no contexto CG, sendo por exemplo responsáveis pela metilação de sequências centroméricas e intimamente ligadas ao estabelecimento de heterocromatina constitutiva (Jeddeloh *et al.*, 1999; Soppe *et al.*, 2002). Ambas estão também implicadas na via de siRNAs heterocromáticos (Lippman *et al.*, 2003). Verificou-se, à semelhança do observado para siRNAs homólogos a sequências centroméricas, que embora em linhas *ddm1* e *met1* a metilação CG nos genes de rDNA seja fortemente reduzida nas regiões homólogas aos siRNAs, o mesmo fenótipo de alteração da metilação não é observado em linhas com alelos nulos de outros membros da via de siRNAs heterocromáticos.

Apesar de não terem sido observadas alterações significativas nos níveis de expressão dos genes ribossomais em consequência da disrupção da via de siRNAs, são no entanto observáveis, ao nível do fenótipo nuclear, alterações na organização da cromatina ribossomal. Todas as mutações em qualquer dos genes da via analisada produzem descondensação do bloco heterocromático característico dos *loci* de rDNA em *A. thaliana*.

Influência da via de siRNAs heterocromáticos no estabelecimento da dominância nucleolar.

Uma vez estabelecida a zona de potencial acção dos siRNAs de rDNA e a sua inclusão na via heterocromática dependente de siRNAs endógenos no diplóide *A. thaliana*, procedeu-se à análise do papel desta via no estabelecimento da dominância nucleolar no alopoliplóide *A. suecica*. As regiões promotoras dos genes de rDNA de *A. thaliana* e *A. arenosa*, embora homólogas no intervalo entre as posições -10 a +30 (tomando +1 como o local de início da transcrição do percursor dos RNAs ribossomais), possuem suficiente polimorfismo nas sequências imediatamente adjacentes para o desenho de sondas de RNA especificas de cada genoma parental. Numa primeira análise, procedeu-se à caracterização dos siRNAs parentais nas linhas LC1 e 9502 do alopoliplóides natural *A. suecica*, Na linha 9502 não se observa dominância nucleolar (Pontes *et al.*, 2003), assim como nas linhas trangénicas de *A. suecica* LC1: DRM2-RNAi, HDT1-RNAi e HDA6-RNAi. Desde logo se tornou evidente que os siRNAs de *A. thaliana* se acumulam a níveis substancialmente reduzidos em comparação com os observados na espécie diplóide. As espécies predominantes de siRNAs de rDNA em *A. suecica* provêm de sequências de rDNA de *A. arenosa*, que curiosamente só são detectados com um sonda de RNA homóloga à regiao -40 a -10 do promotor de rDNA, não se sobrepondo ao local de início de transcrição (+1). Tanto em *A. suecica* 9502 como em *As*DRM2-RNAi observa-se uma redução nos níveis de siRNAs, indiciando o envolvimento do mecanismo de siRNA heterocromáticos na dominância nucleolar. Por outro lado, *As*HDT1-RNAi não mostra qualquer efeito na biogénese dos siRNAs e em *As*HDA6-RNAi observa-se mesmo uma maior acumulação de siRNAs de ambos os progenitores, levantando a hipótese de a via dos siRNAs não será o seu único factor determinante, embora importante para o estabelecimento da dominância nucleolar,.

De forma a melhor compreender a função deste mecanismo epigenético mediado por RNA, procedeu-se ao "knock down" em *A. suecica* LC1 dos genes *NRPD2a*, *DCL3*, *AGO4*, *RDR2* e *HEN1*. Verificou-se que em todos os casos foi possível associar decréscimos de acumulação de transcritos dos genes alvo à redução de metilação do DNA de uma sequência endógena controle (*AtSN1*) e à redução dos níveis de acumulação de siRNAs. A análise dos transcritos de rDNA provenientes de cada genoma parental demonstrou o envolvimento desta via no estabelecimento da dominância nucleolar, uma vez que foram identificados transcritos provenientes de genes de rDNA do complemento

genómico de *A. thaliana*. A reactivação dos genes de rDNA de *A. thaliana* é acompanhada por alterações ao nível da organização da cromatina ribossomal no núcleo poliplóide, excepto em plantas transgénicas *As*AGO4-RNAi. Esta alteração no fenótipo nuclear, passando os *loci* de rDNA de um ou dois blocos condensado na periferia nucleolar característicos em *A. suecica* LC1, para uma organização mais descondensada, encontra-se associada à reactivação dos genes ribossomais de *A. thaliana* (Pontes *et al.*, 2003; Lawrence *et al.*, 2004; Earley *et al.*, 2006).

Em conclusão, os resultados obtidos no decurso do trabalho apresentado demonstram o papel da via de siRNAs heterocromáticos na modulação da cromatina ribossomal em *Arabidopsis sp.*. Embora não se tenham identificado alterações ao nível da metilação de sequências de DNA ou analisado as marcas pós-tradução de histonas, foi estabelecida uma correlação directa entre siRNAs homólogos à região promotora dos genes de rDNA e a detecção de alterações ao nível dos padrões de organização da cromatina ribossomal em núcleos interfásicos. A reactivação dos genes de rDNA do complemento genómico de *A. thaliana* no alopoliplóide *A. suecica* demonstra a importância da via de siRNAs heterocromáticos no estabelecimento da dominância nucleolar, abrindo caminho a estudos futuros que nos poderão aproximar da compreensão global do controlo deste fenómeno epigenético.

Acknowledgments

The tortuous path leading to graduation although, to some extent, a solitary journey would not have been possible without the help and support of both supervisors, colleagues, friends and family. It is with great pleasure that I acknowledge and thank the contribution of all of those directly or indirectly involved in this work, without which it would have just not been the same.

Table of Contents

I - Introduction

I.1 The eukaryotic nucleus

Eukaryotes store their genetic information in a specialized cellular structure, the nucleus. In this cellular compartment, DNA molecules are associated with structural proteins, especially histones, to form chromatin and higher order structures that overcome the spatial constraints in the nucleus (Ridgway and Almouzni, 2001). The basic chromatin unit is the nucleosome, consisting of 147 base pairs (bp) of DNA tightly wrapped 1.7 times around a protein octamer composed of two each of histones H2a, H2b, H3 and H4. Adjacent nucleosomes are connected by 30-40 base pairs of linker DNA (Luger *et al.*, 1997). The nucleosome array, which resembles as "beads on a string" when viewed by electron microscopy is further folded with the participation of histone H1 into a 30 nm chromatin fiber. At this level, DNA is ~30 to 40 folds more compact than in its linear dimension, with further levels of chromatin folding leading to poorly understood organizations into large chromatin domains and chromosome territories (Belmont *et al.*, 1999).

Chromatin is a dynamic structure that can be modulated to correspond to the cell's demands at a given time for a particular set of genes. The organization of DNA and the histone octamer into nucleosomes and the regulation of the higher-order condensation of chromatin not only play a role in transcriptional activation and repression, but are required for stable silencing and differentiation (Francis and Kingston, 2001; Horn and Peterson, 2002). This plasticity enables responses to environmental stress or to adapt the transcriptome to developmental needs, ensuring that a cell maximizes its resources to efficiently respond to its particular context. Chromatin organization is also important in fundamental nuclear processes including replication, recombination and repair. Chromatin can adopt different structural and functional states tightly linked to changes in gene activity, related to changes in the folding of the nucleosome array (Eberharter and Becker, 2002) as well as to changes in covalent modification of the core histone proteins (Strahl and Allis, 2000).

In recent years a new mechanism for chromatin modulation has been identified, and with it the notion that chromatin modifications can be targeted with high specificity. The major player is RNA and this realization has revolutionized the way we interpret transcriptional and post-transcriptional silencing, stress responses, development and chromatin modulation.

I.1.1. Heterochromation and euchromatin

The first reference to the different properties of chromatin states termed heterochromatin and euchromatin was made by Heitz (1928). The author observed differential staining of chromosomes in interphase nuclei, revealing segments that remained condensed at all times and displayed a more intense staining, which he referred to as positive heteropycnosis. The differential staining led to the notion that chromosomes are organized in two distinct chromatin environments, euchromatin and heterochromatin. The term heterochromatin refers to highly packaged DNA that remains condensed throughout the cell cycle in a structure relatively inaccessible to DNA-modifying factors. By contrast, euchromatin is decondensed during interphase and tends to be transcriptionally active (Cremer and Cremer, 2001). Hetero- and euchromatin differ in gene density, content of repetitive DNA sequences, chromatin composition, nucleosome spacing and accessibility to nucleases (Henikoff, 2000).

Typical heterochromatin is late replicating during S phase (Henikoff, 2000; Karpen and Allshire, 1997) and can act as a matchmaker in recognition, pairing and proper segregation of homologous chromosomes at meiosis (Dernburg *et al.*, 1996). Protein components of heterochromatin, such as HP1 (**H**eterochromatin **P**rotein **1**) and related proteins were shown to be essential for correct chromosome segregation (Bernard *et al.*, 2001; Eissenberg and Elgin, 2000; Ekwall *et al.*, 1995; Kellum and Alberts, 1995; Nonaka *et al.*, 2001). Heterochromatin is a feature of constitutively condensed and largely inactive telomeres, centromeres and pericentromeric regions, which therefore compose what is known as constitutive heterochromatin (Henikoff, 2000). Facultative heterochromatin defines domains that are transcriptionally inactive in certain cell lineages or developmental stages but active in others (Grewal and Elgin, 2002), being the silencing of euchromatic regions due to chromatin structural changes leading to its heterochromatinization (Breszki and Jerzmanowski, 2003; 2004). Most importantly, heterochromatin can be transmitted through mitosis, thus forming a basis for epigenetic inheritance and cellular memory. Increasing evidence indicates that the initiation and maintenance of silenced heterochromatin involves repetitive sequences (Fourel *et al.*, 2002), the abundant chromatin histone proteins and RNA interference pathways that will be discussed in more detail below (Hannon, 2002; Lippman *et al.*, 2003; Onodera *et al.*, 2005).

I.1.2. Histone post-translational modifications mark chromatin states

Histones have been conserved during evolution. However, in contrast to the universal nature of DNA, they are dynamically altered by post-translational modifications. The central part of a histone molecule is a globular domain formed by three helices, also known as the histone fold (Luger *et al.*, 1997). Histone N-terminal 'tails' contain large numbers of the basic amino acids lysine and arginine, resulting in a positive net charge of the tail at physiological pH, and mediate internucleosomal contacts

(Luger *et al.*, 1997). Amino acid residues at specific positions within histone tails are subject to a number of post-translational modifications, which cause structural and functional rearrangements in chromatin and define what has been called the 'epigenetic histone code' (Jenuwein and Allis, 2001; Turner, 2002). The histone code is thought to dictate nucleosomal interactions and the association of proteins that collectively influence chromatin packaging and gene regulation (Rice and Allis, 2001; Strahl and Allis, 2000; Turner, 2000). The modification states of the N-terminal tails of histones H3 and H4 appear to play a major role in heterochromatin formation (Richards and Elgin, 2002).

I.1.2.1. Histone Acetylation

The reversible acetylation of N-terminal lysine residues at positions 5, 8, 12, and 16 of histone H4 and positions 9, 14, 18, and 23 of histone H3 mediates decondensation of the nucleosome structure (Garcia-Ramirez *et al.*, 1995; Loidl, 1994), alters histone–DNA interactions (Hong *et al.*, 1993), and facilitates access and binding of transcription factors to genes transcribed by RNA polymerases II or III (Vettese-Dadey *et al.*, 1996). Histone acetylation is correlated with transcriptional activity (Csordas, 1990; Grunstein, 1997; Loidl, 1994; Struhl, 1998; Turner *et al.*, 1992) and deacetylation is associated with transcriptional silencing (Grunstein, 1997; 1998; Spencer and Davie, 1999). Accordingly, transcriptionally silenced heterochromatic domains are usually less acetylated than euchromatin. Histone hypoacetylation and heterochromatin assembly can be mediated by the Sir-protein complex in a step-wise process (Hall *et al.*, 2002; Grewal and Jia, 2007; Grunstein, 1998; Volpe *et al.*, 2002). In female mammals, the inactivated X chromosome is organized into facultative heterochromatin and is largely free of acetylated histones (Jeppesen and Turner, 1993; Heard, 2005).

The histone acetylation level on chromatin results from a balance between histone acetyltransferase (HAT) and histone deacetylases (HDAC) activities (Vogelauer *et al.*, 2000). HATs and HDACs basically fulfill similar functions in all eukaryotes, forming a large family for which many members have been shown to be responsible for the reversible and dynamic changes of histone tails (Loidl, 2004). In general, HDACs are correlated with transcriptional repression and gene silencing, cooperating with histone methyltransferases (HMTs) and DNA methyltransferases (Pikaard and Lawrence, 2002). Mutations in the gene encoding the *Arabidopsis* HDAC HDA6, revealed its involvement in transgene silencing (Aufsatz *et al.*, 2002; Murfett *et al.*, 2001) and HDAC function is implicated in the global control of gene regulation in plants (Gao *et al.*, 2003; Pipal *et al.*, 2003; Rossi *et al.*, 2003; Wu *et al.*, 2000).

I.1.2.2. Histone Methylation

Evidence for histone methylation was first demonstrated by Murray (1964). Histone methylation occurs on arginine (R2, R17, R26 of H3, and R3 of H4) and lysine residues (K4, K9, K27 and K79 of H3, also K20 of H4 in mammals). Mono- or dimethylated arginines and mono-, di-or trimethylated lysines have been reported (Bannister et al., 2002), and are correlated with transcriptional regulation (Zhang and Reinberg, 2001). Functional differences arise from different histone methylation modifications (Dutnall, 2003). In mammalian cells, trimethylated H3K9 is preferentially localized to pericentromeric heterochromatin, di- and trimethylated to facultative heterochromatin while monomethylated and dimethylated H3K9 are localized to euchromatin (Peters *et al.*, 2003; Rice *et al.*, 2003). In *Neurospora*, silent *loci* are also associated with H3^{trimethyl}K9 (Tamaru *et al.*, 2003). In *Arabidopsis*, both monomethylated and dimethylated H3K9 are found at silent *loci*, but H3^{trimethyl}K9 has not been detected (Jackson *et al.*, 2004).

Rea *et al.* (2000) have shown that mammalian homologues of *Drosophila* SUPPRESSOR OF VARIEGATION39 (SUV39), encodes H3K9-specific methyltransferases (Lachner and Jenuwein, 2002). The methyltransferase activity of SUV39 is directed against lysine 9 of histone H3 and its catalytic domain resides within a highly conserved structure, the SET domain, common to a protein family that can influence gene expression through histone methylation. Methylated H3K9 recruits the highly conserved protein HP1 (Bannister *et al.*, 2001; Jackson *et al.*, 2002; Lachner *et al.*, 2001) and is thereby involved in heterochromatin assembly (Nakayama *et al.*, 2001; Zhang and Reinberg, 2001). On the other hand, lysine-specific demethylase activity of histones has been recently reported in humans. LSD1 is required for demethylation of both H3K4 and H3K9 from dimethylated to an unmethylated form. Its specificity for each of the post-translational modifications, and hence its ability to influence silencing or transcriptional activation, is dependent on interaction with protein cofactors (Lee *et al.*, 2005; Metzger *et al.*, 2005; Shi *et al.*, 2004; Shi *et al.*, 2005). Additionally, H3^{trimethyl}K9 is the substrate for the action of the demethylase JMJD2b (Fodor *et al.*, 2006).

I.1.3. DNA methylation

Cytosine methylation on carbon 5 of the pyrimidine ring is the most common DNA modification in eukaryotes, playing an important role in gene regulation and genome stability. DNA methylation is associated with processes such as X chromosome inactivation and imprinting in vertebrates and immobilization of transposable elements in higher plants (Finnegan *et al.*, 1998; 2000). DNA methylation is also associated with silencing protein-coding genes in plant polyploids (Lee and Chen, 2001) and regulating excess RNA polymerase I-transcribed ribosomal RNA genes (Chen and Pikaard, 1997a).

The distribution of methylcytosine is not random; most methylated residues occur within repetitive DNA found in heterochromatin (Bennetzen, 1996; Bennetzen *et al.*, 1998; Guseinov *et al.*, 1975). However, methylcytosine is also found in single-copy genes where it is implicated in regulating gene expression and its distribution is linked to RNA polymerase activity (Zilberman *et al.*, 2007). Methylcytosine can occur in any sequence context in plant DNA (Meyer *et al.*, 1994) but is most common in the CpG and CpNpG symmetric contexts (Gruenbaum *et al.*, 1981; Zilberman *et al.*, 2007; Zhang *et al.*, 2006).

The methylation of DNA strands requires specific DNA methylases. *Arabidopsis* DNA METHYLTRANSFERASE1 (MET1), the homolog of mammalian DNMT1, is responsible for the maintenance of DNA methylation at CpG sites (Cao *et al.*, 2000; Goodrich *et al.*, 1997; Lehnertz *et al.*, 2003; Saze *et al.*, 2003), which forms the majority of methylated sites, but is also needed for CpNpG methylation to a lesser extent. Hypomethylated *met1* plants display a wide range of pleiotropic effects (Finnegan *et al.*, 1998; 2000; Finnegan and Kovak, 2000; Goodrich *et al.*, 1997), which are similar to those of *Arabidopsis* plants with suppressed histone deacetylase activity. CHROMOMETHYLASE3 (CMT3), which is unique to plants (Meyerowitz, 2002), maintains mostly methylation of CpNpG with a lesser effect on asymmetric (CNN) sites (Birve *et al.*, 2001; Yadegari *et al.*, 2000). *cmt3* mutations affect methylation of a subgroup of genes, such as *SUPERMAN* or *PHOSPHORIBOSYLANTHRANILATE-2* (*PAI2*), and retrotransposons (Birve *et al.*, 2001; Yadegari *et al.*, 2000). DOMAINS-REARRANGED METHYLASE2 (DRM2 - the homolog of the human DNMT3 methyltransferases) acts as a *de novo* methyltransferase in genome defense (Cao and Jacobsen, 2002) and has partially overlapping and redundant function with CMT3 with respect to CpNpG methylation (Bartee *et al.*, 2001; Lindroth *et al.*, 2001). It is intimately associated with RNAi in the establishment of repressive epigenetic marks at small-RNA targeted DNA regions in plants (Cao *et al.*, 2003; Chan *et al.*, 2004).

To counteract DNA methyltransferases plants rely on the activity of DNA demethylases and therefore are not dependent on DNA replication to remove methylation. DEMETER and **R**EPRESSOR **O**F **S**ILENCING **1** (ROS1) are known examples of DNA glycosylases involved in demethylating DNA(Agius *et al.* 2006; Morales-Ruiz *et al.*, 2006) and might confer plasticity to the genome in responding to biotic and abiotic stresses.

I.1.4 The Epigenome and coordination of epigenetic marks

The control of gene transcriptional activity, and the establishment of euchromatin or heterochromatin, results from the coordination of various epigenetic marks. Histone deacetylation and DNA methylation are intimately associated with gene repression (Richards and Elgin, 2002). In

animals, the two processes are mechanistically linked via proteins that contain methylcytosine-binding domains (MBD), such as methyl-CpG-binding protein (MeCP2) and MBD1-4. These proteins bind to methylated DNA and recruit HDACs (Ogas *et al.*, 2000). The epigenetic information may also flow from the histone to the cytosine methylation system. The HDAC inhibitor TSA leads to cytosine hypomethylation in *Neurospora* (Selker, 1998), and a similar effect has been noted in mammalian cells (Cervoni and Szyf, 2001) and plants (Lawrence *et al.*, 2004). Silencing of one X chromosome in mammalian females is intimately associated with histone deacetylation, H3K9 and DNA methylation in a process triggered by long non-coding RNAs which act in *cis* (Heard, 2005; Penny *et al.*, 1996). *S. pombe*, which is unable to methylate DNA, relies on its RNAi machinery to induce H3K9 methylation and histone deacetylation at centromeric sequences, thus inducing heterochromatin formation (Grewal and Jia, 2007). Furthermore, deacetylation of specific lysines such as H3K9 is a prerequisite for H3K9 methylation and subsequent recruitment of proteins (HP1) that are involved in heterochromatin formation (Lachner and Jenuwein, 2002; Rice and Allis, 2001). Some histone methyltransferases contain predicted methylcytosine-binding domains (MBDs), which indicate that DNA methylation can also direct histone methylation (Zhang and Reinberg, 2001).

I.2. *RNA mediated post-transcriptional and epigenetic silencing mechanisms*

RNA based silencing mechanisms are diverse yet highly conserved phenomena that mediate gene silencing in a sequence-specific manner. RNA interference (RNAi) was first described in the nematode *Caenorhabditis elegans* (Fire *et al.*, 1998) and in transgenic plants (Baulcombe, 2004), and is known to be a posttranscriptional gene silencing (PTGS) mechanism triggered by double-stranded RNAs (dsRNAs). Since the first discovery of RNAi our knowledge of RNA mediated silencing pathways has increased immensely and is now known to comprise a set of highly diversified pathways in multicellular organisms.

Small RNAs (smRNAs) are short single-stranded RNA molecules with sizes ranging between 21 and 28 nucleotides (nt). They can be found in organisms as diverse as fungi (*Schizosaccharomyces pombe*), protists (*Tetrahymena thermophila*), plants (*Arabidopsis sp.*) and animals (*Drosophila sp.* and mammals) and are known to be involved in transcriptional (TGS) and post-transcriptional gene silencing (PTGS) pathways resulting in such processes as heterochromatin assembly (fission yeast, plants flies, mammals) programmed DNA elimination (*Tetrahymena*) or translational arrest (worms, flies, plants, animals). In these ways, RNA silencing mechanisms function in host genome defense against viruses and transposable elements, and in diverse cellular processes such as spatial and temporal developmental regulation or regulation of chromosome structure (Bartel, 2004; Baulcombe, 2004; Bernstein and Allis, 2005; Grewal and Jia, 2007; Hannon, 2002; Waterhouse *et al.*, 2004).

Central to smRNA function is a role in guiding effector complexes to both DNA and/or RNA targets in a homology dependent manner, hence conferring specificity to the various mechanisms in which smRNAs are involved (Ambros, 2004; Bartel, 2004; He and Sontheimer, 2004; Kidner and Martienssen, 2004; 2005).

Whole genome approaches have characterized the distribution and abundance of smRNAs in *Arabidopsis thaliana* (Llave *et al.*, 2002; Lu *et al.*, 2005). The major sources for siRNAs (small interfering RNAs of RNAi-related pathways) are intergenic regions, transposable and retro-transposable elements, centromeric regions and rDNA. In contrast, few smRNA originate from protein coding genes. A significant study was conducted in *S. pombe* where the distribution of smRNAs correlated closely with the distribution of histone H3 methylated at lysine 9, revealing a mechanism whereby smRNAs can direct heterochromatic modifications to specific chromosomal regions (Cam *et al.*, 2005).

I.2.1. Small interfering RNAs and Micro RNAs

Small RNA biogenesis is initiated from dsRNA precursors. Depending on the origin and length of these precursors distinct types of small RNA classes have been defined in RNA-mediated silencing pathways. The best-known forms are micro RNAs (miRNAs) and small interfering RNAs (siRNAs), which are produced by similar pathways and have similar modes of action (Ambros, 2004; Bartel, 2004; He *et al.*, 2005). The former are produced through processing of short endogenous hairpin precursor transcripts, targeting other *loci* with similar but not identical sequences for translational repression or target perfectly homologous mRNAs for degradation (Yekta *et al.*, 2004). By contrast siRNAs are usually produced from longer double-stranded RNAs, such as viral replication intermediates, bidirectionally transcribed endogenous repetitive sequences, or RNAs whose complementary strands are generated by an RNA-dependent RNA polymerase. In addition to roles in PTGS, siRNAs can also direct transcriptional gene silencing, acting at the genome level to direct epigenetic modifications, namely DNA and histone methylation, to homologous DNA sequences (Volpe *et al.*, 2002; 2003; Hall *et al.*, 2003).

Another distinguishable feature in the biogenesis of the two smRNA classes is the number of distinct smRNA molecules produced from the precursor dsRNA. Typically, processing of miRNA hairpin precursors gives rise to a single and defined miRNA molecule whereas siRNA dsRNA precursor processing produces multiple species in both sense and antisense orientation, sometimes in a perfect tandem arrangement (Allen *et al.*, 2005; Gasciolli *et al.*, 2005; Vasquez *et al.*, 2004). Stability of these smRNA molecules in *Arabidopsis* is dependent on the methylation of the 3' terminal nucleotide (conversion of 2'-OH to 2'-O-CH$_3$) by the action of HEN1 (*HUA ENHANCER1*);

accumulation levels of smRNAs in *hen1* mutants are substantially reduced. The HEN1 methylation activity, exerted on smRNA duplexes prior to unwinding into the functional single stranded molecules, was shown to protect smRNAs duplexes from uridylation (Li *et al.*, 2005), which is speculated to trigger smRNA degradation. 3' end methylation may also be an important determinant of smRNA stability in mammals (Amarzguioui *et al.*, 2003).

I.2.2. Dicer enzymes

Although some variations have been found among organisms, from plants to animals, RNA mediated silencing is typically initiated by the processing of dsRNAs by Dicer RNase-III-like endonucleases. DsRNA precursors are converted into ~21-28 bp molecules that contain 2nt 3' end overhangs with free hydroxyl groups and are phosphorylated at both 5' ends (Beirnstein *et al.*, 2001; Carmell and Hannon, 2004; Kurihara and Watanabe, 2004). Dicer enzymes are complex multifunctional proteins containing an RNA helicase domain, a PAZ RNA binding domain as well as the canonical RNase III-like ribonuclease and dsRNA binding motifs (Schauer *et al.*, 2002; Song *et al.*, 2003). Diverse Dicer-like (DCL) enzymes have been characterized, suggesting multiple activities. Many animals, including *Caenorhabditis elegans*, have a single Dicer enzyme that mediates all the required processing of dsRNAs into both siRNA and miRNAs (Grishok *et al.*, 2001; Ketting *et al.*, 2001; Knight and Bass, 2001). In contrast, there are two Dicer paralogues in *Drosophila melanogaster* where one is responsible for miRNA biogenesis and the other for siRNA production (Lee *et al.*, 2004; Pham *et al.*, 2004; Tjisterman and Plasterk, 2004). In the filamentous fungus *Neurospora crassa* two Dicer proteins are functionally redundant in siRNA biogenesis (Catanalotto et al., 2004). The *A. thaliana* genome encodes four DCL enzymes (Schauer *et al.*, 2002) which participate in distinct smRNA pathways, with some pathways requiring the activity of more than one DICER. DICER-LIKE 1 (DCL1) (Papp *et al.*, 2003) is involved in the processing of miRNA precursors (Bartell, 2004). DCL2 is required for producing 22nt viral small interfering RNAs and natural antisense siRNAs (nat-siRNAs) (Borsani *et al.*, 2005; Xie *et al.*, 2004). DCL4 is needed in the generation of 21nt trans-acting siRNAs (ta-siRNAs), which are siRNAs that regulate, in *trans*, endogenous genes by targeting homologous mRNAs in a manner very similar to miRNAs (Gasciolli *et al.*, 2005). This pathway is dependent on miRNAs, and hence on the activity of DCL1, to set the phase for ta-siRNA biogenesis. Finally, DCL3 (Xie *et al.*, 2004) is exclusively involved in the production of the 24nt siRNA class specifically involved in RNA mediated chromatin modifications by directing RISC-like complexes (**R**NA-**I**nduced **S**ilencing **C**omplexes) to homologous DNA regions, in a pathway termed **R**NA-**d**ependent **D**NA **M**ethylation (RdDM) (Mette *et al.*, 2002).

I.2.3. The Argonaute protein family

The smRNAs generated by DICER activity are loaded directly into Argonaute proteins (AGO) that form the core of active RISC complexes (Chendrimada *et al.*, 2005; Hammond *et al.*, 2001; Tabara *et al.*, 2002). These complexes then carry out the identification and silencing of the target sequence matching the smRNA (Caudy *et al.*, 2002; Hannon, 2002; Silva *et al.*, 2002). In mammals, AGO2 is the slicer protein of RISC that is directly responsible for mRNA cleavage (Liu *et al.*, 2004; Song *et al.*, 2004). Through genetic, biochemical and crystal structure analyses, it was shown that the AGO2 PIWI domain is closely related to RNase H family members and transposases, allowing AGO proteins to cleave their RNA targets in smRNA/mRNA hybrids (Carmell *et al.*, 2002; Song *et al.*, 2004). Further cytological analysis demonstrated that AGO-containing RISC mRNA cleavage takes place in cytoplasmic P-bodies (Liu *et al.*, 2005a; 2005b). *C. elegans* contains 27 AGO genes that have been proposed to act in two distinct steps of smRNA biogenesis. Translation arrest mediated by miRNAs is thus far restricted to the action of ALG-1 and ALG-2 (Grishok and Mello, 2002), but response to both endogenous and exogenous RNA triggers involves a more complex pathway. Primary siRNAs are loaded into ERGO-1 and RDE-1 (Tabara *et al.*, 1999), respectively, which target homologous mRNAs for cleavage and create RNA templates for secondary siRNA biogenesis. These secondary siRNAs direct gene silencing through the redundant action of at least 3 other AGO proteins that are apparently depleted of slicer activity (Yigit *et al.*, 2006).

In *Arabidopsis sp.*, ten Argonaute proteins have been identified, suggesting evolutionary specialization of these proteins to non-redundant roles in RNA silencing mechanisms (Fagard *et al.*, 2000; Morel *et al.*, 2002). However, members of the Argonaute family in *Arabidopsis* other than AGO1 and AGO4 lack functional characterization. AGO7 has been implicated in the ta-siRNA pathway, specifically in PTGS mediated by the TAS3 locus, as *ago7* mutants display abnormal leaf development (Adenot *et al.*, 2006). AGO1 was isolated through genetic screens and is required for posttranscriptional gene silencing (Bohmert *et al.*, 1998; Fagard *et al.*, 2000). Biochemical analysis showed that AGO1 co-immunoprecipitates with miRNAs and ta-siRNAs (Baumberger and Baulcombe, 2005; Qi *et al.*, 2006), which are small RNAs involved in mRNA target cleavage. Furthermore, in *ago1* null mutants, decreased miRNA levels are accompanied by target mRNA accumulation (Vaucheret *et al.*, 2004). AGO1 was the first Argonaute protein shown to have slicing activity (Baumberger and Baulcombe, 2005; Qi and Hannon, 2005). In addition to its contribution to mRNA cleavage, AGO1 and another family member, AGO4, are involved in gene expression regulation by chromatin modification (Chan *et al.*, 2004; Lippman *et al.*, 2003; Lippman *et al.*, 2004; Morel *et al.*, 2002; Zilberman *et al.*, 2003; Zilberman *et al.*, 2004). AGO4 is required for 24nt siRNA

production and DNA methylation of some endogenous *loci*, but does not affect miRNA production (Chan *et al.*, 2004; Lippman *et al.*, 2003; Lippman *et al.*, 2004; Zilberman *et al.*, 2003; Zilberman *et al.*, 2004). This Argonaute was shown to co-immunoprecipitate with 23-24nt siRNAs (Li *et al.*, 2006; Qi *et al.*, 2006) and also displays slicer activity (Qi *et al.*, 2006); a function that does not impair its siRNA mediated targeting function although in some target sequences downstream DNA methylation is somewhat impaired. It has been proposed that AGO4 slicer activity might be important for generation of secondary siRNAs, following the same principle involved in ta-siRNAs (Gasciolli *et al.*, 2005) or siRNAs in S. pombe and *C. elegans* (Yigit *et al.*, 2006), in order to reinforce the silencing signal. AGO6 has been recently implicated in RdDM, playing a partially redundant role with AGO4 at some target *loci* (Zheng *et al.*, 2007).Theoretically, seven other AGO proteins are predicted to have the catalytic active site of a slicer (Qi *et al.*, 2005; Sigova *et al.*, 2004) and the determination of their function is likely to render new insights to the RNAi pathways in *Arabidopsis*.

I.2.4. RNA dependent RNA polymerases

Some organisms have genes for RNA-dependent RNA polymerases (RdRP), which can synthesize dsRNA from single stranded RNAs and can play roles in the initiation and/or amplification of siRNA production (Baulcombe, 2004; Makeyev and Bamford, 2002; Schielbel *et al.*, 1993). Genome sequence data analysis has revealed that *C. elegans*, *N. crassa* and *A. thaliana* contain several RdRPs, but no equivalent proteins are found in *D. melanogaster* or mammals (Maine, 2001; Schwarz *et al.*, 2002). *Rdp1* in *S. pombe* encodes an RdRP involved in the amplification of the siRNA signal required for centromeric heterochromatin establishment (Motamedi *et al.*, 2004). *RRF-1* (Smardon *et al.*, 2000) and *RRF-3* (Simmer *et al.*, 2002) are required for PTGS in *C.* elegans, and are also involved in secondary siRNA biogenesis in endogenous or exogenous RNAi pathways, respectively, following slicing by the primary Argonautes RDE-1 and ERGO-1. *A. thaliana* has six predicted RdRPs but thus far only *RDR6* (Mourrin *et al,* 2000) and *RDR2* (Xie *et al.*, 2004) have determined functions. RDR6 is required for viral and transgene silencing (Allen *et al.*, 2005; Balcoumbe, 2004; Dalmay *et al.*, 2000; Mourrin *et al,* 2000; Peragine *et al.*, 2004; Xie *et al.*, 2004) as well as for biogenesis of ta-siRNAs (Adenot *et al.*, 2006; Gasciolli *et al.*, 2005) and nat-siRNAs (Borsani et *al.*, 2005). Similar to what is observed in *S. pombe* and *C. elegans*, RDR6 might be responsible for generation of secondary RNAs, thus amplifying the silencing signal, if one considers the biogenesis of ta-siRNAs as an RDR6 trademark. In this case, after targeted cleavage of a non-coding RNA by a miRNA loaded RISC, RDR6 is involved in the generation of secondary dsRNA precursors that will act as substrates for DCL4 processing into 21nt smRNA duplexes (Axtell *et al.*, 2006; Gasciolli *et al.*, 2005). This RdRP is also involved in developmental processes; *rdr6* mutant lines undergo precocious transition from

juvenile to adult stage and link the RNA mediated PTGS pathway to the temporal control of shoot development in plants (Peragine *et al.*, 2004). siRNA-dependent chromatin modification requires the specialized role of RDR2 (Xie *et al.*, 2004). This RdRP has thus far been exclusively linked to RdDM, together with DCL3. Heterochromatic siRNAs (24nt class) are essentially eliminated in *rdr2* mutants, resulting in transcriptional up-regulation, depletion of DNA methylation and enrichment of euchromatic post-translational histone modifications in target genes (Herr *et al.*, 2005; Xie *et al.*, 2004).

I.2.5. siRNA-induced heterochromatin formation in *S. pombe*

The link between siRNA-specific targeting and methylated H3K9 has been studied in great detail in *Schizosaccharomyces pombe*, the system in which RNA-mediated heterochromatin formation was first described (Grewal and Rice, 2004; Volpe *et al.*, 2002). In *S. pombe*, *dcr1* (Dicer homolog) and *ago1* (Argonaute homolog) mutations were shown to reduce centromeric-repeat H3K9 methylation, which is necessary to maintain centromere function (Volpe *et al.*, 2002). DCR1-processed siRNAs corresponding to centromeric repeats interact with AGO1, the chromodomain protein CHP1 and an uncharacterized protein, TAS1, to form the RNA-induced initiator of transcriptional gene silencing (RITS), which guides siRNAs to complementary sites on the genome (Verdel *et al.*, 2004). The RITS complex, and specifically AGO1 and CHP1, then recruits the chromatin-modifying factors Swi6, an HP1 homolog, and the histone methyltransferase CLR4 in order to silence target regions. The presence of siRNAs in RITS also requires an RNA-directed RNA polymerase complex (RDRC), which consists of the *S. pombe* RdRP, RDRP1, the putative helicase HRR1 (helicase required for RNA-mediated heterochromatin assembly 1) and CID12, a 38-kDa protein involved in mRNA polyadenylation (Motamedi *et al.*, 2004). The RDRC complex is likely critical not only for RITS stability but also for triggering secondary dsRNAs production and in this way amplifying siRNAs and the silencing signal (Motamedi *et al.*, 2004). SiRNA heterochromatin formation in *S. pombe* also requires RNA polymerase II activity (pol II), suggesting a direct link between target locus transcription and siRNA pathway initiation (Kato *et al.*, 2005; Schramke *et al.*, 2005). However it remains elusive how pol II transcription is coupled to siRNA production to guide chromatin modifications. Two models have been proposed: direct RNA-DNA pairing (DNA-recognition model) or pairing of siRNAs and a nascent transcript from the target locus (RNA-RNA recognition model) (Grewal and Moazed, 2003; Matzke and Birchler, 2005). In the first model, a transcript homologous to the target gene remains associated with the gene, acting as a local address to allow siRNA guided chromatin modification complexes, access the target gene. Alternatively in the siRNA/DNA model, RNA pol II might unwind the target DNA enough to allow siRNAs and the RNAi machinery access to the target promoter.

I.2.6. The plant specific RNA-dependent DNA methylation pathway

In contrast to *S. pombe*, in which there is no detectable DNA methylation, plant RNA silencing is frequently accompanied by DNA cytosine methylation, usually coupled to H3K9 methylation (Jackson *et al.*, 2002; Malagnac *et al.*, 2002; Soppe *et al.*, 2002). Through RdDM, 24nts siRNAs lead to *de novo* methylation of cytosines in all sequence contexts within a region of RNA-DNA sequence homology at endogenous loci.

This mechanism requires the presence of RDR2, which presumably produces dsRNAs from single stranded templates, dicing of 24nt siRNAs by DCL3, and AGO4 as the siRNA-binding protein that presumably targets corresponding *loci* for heterochromatic modifications (Cao and Jacobsen, 2002; Cao *et al.*, 2003; Xie *et al.*, 2004). Target-site establishment of *de novo* DNA methylation requires DRM2, which can methylate cytosines in any sequence context in a process still not fully understood (Cao *et al.*, 2003). Surprisingly, the maintenance cytosine methyltransferase1, MET1 and DDM1 (DECREASE IN DNA METHYLATION 1), a SWI2/SNF2-family chromatin-remodeling factor are also required for siRNA production at certain loci (Lippman *et al.*, 2003; Zilberman *et al.*, 2004). Therefore, siRNA targeting of epigenetic modifications sustains siRNA production, likely as part of a positive feedback loop reinforcing silencing through continuous siRNAs biogenesis. In turn, MET1 can be responsible for the inheritance of a silent heterochromatic state in the absence of siRNAs through maintenance methylation, mostly at CpG sites, but also at CpNpG sites, at each round of DNA replication (Aufsatz *et al.*, 2004; Cao *et al.*, 2003).

A recently discovered activity required for siRNA-directed DNA methylation is plant-specific nuclear RNA polymerase IV (pol IV), (Herr *et al.*, 2005; Onodera *et al.*, 2005; Pontier *et al.*, 2005; The Arabidopsis genome Initiative, 2000; this study). Reverse and forward genetic studies identified two pol IV largest subunits, designated NRPD1a and NRPD1b, and two potential second-largest subunits, NRPD2a and NRPD2b (Herr *et al.*, 2005; Kato *et al.*, 2005; Onodera *et al.*, 2005). However, only the second largest subunit encoded by *NRPD2a* is transcribed, with *NRPD2b* being a non-functional allele (Onodera *et al.*, 2005). NRPD1a and NRPD2a are both required for siRNA accumulation and DNA methylation at heterochromatic loci (Herr *et al.*, 2005; Onodera *et al.*, 2005). By contrast, an *nrpd1b* mutant allele leads to loss of DNA methylation, but this subunit is not essential for siRNA production (Kanno et al., 2005). Therefore, the two functional forms of pol IV, pol IVa and pol IVb, both of which use NRPD2a as their second-largest subunit, appear to play distinct roles in RNA silencing. Pol IVa is upstream in the pathway, likely providing transcripts used as substrates by RDR2, whereas pol IVb transcription might play a role in guiding RNA-dependent DNA methylation (or histone modification) complexes to chromatin targets (Herr *et al.*, 2005; Kanno *et al.*, 2005; Pontes

et al. 2006). The precise roles of RNA polymerase IVa and IVb in siRNA-mediated chromatin modifying pathways remains largely unclear but several structural differences suggest atypical functions for pol IV. Although pol IV largest and second-largest subunits are similar to catalytic subunits of RNA polymerases I, II and III, in vitro assays failed to detected a DNA-dependent RNA polymerase activity (Onodera *et al.*, 2005). Moreover, the active sites in NRPD1a and NRPD1b suggest a modified RNA polymerase activity, possibly using methylated DNA or dsRNA as a template (Herr *et al.*, 2005; Onodera *et al.*, 2005). In addition, NRPD1b has a large C-terminal domain (CTD) reminiscent of the one found in pol II largest subunit. In the latter, the CTD domain is a binding site for other proteins and a target for phosphorylation, helping promote pol II transition from a pre-initiation to an elongation state (Dahmus, 1996; Hampsey and Reinberg, 2003). Recently, it was found that a stable interaction occurs between pol IVb and AGO4, via the NRPD1b CTD-domain, strongly suggesting that a pol IVb-RISC complex might be important for *locus* targeting (Li *et al.*, 2006).

SiRNA-directed DNA methylation also requires the activity of the SWI2/SNF2-like chromatin remodeling protein DRD1 (DEFECTIVE IN RNA-DIRECTED DNA METHYLATION) that, like pol IVb, is not required for siRNA biogenesis but is necessary for *de novo* DNA methylation of target sequences in CpG and non-CpG contexts. DRD1 might function by opening the chromatin structure to make DNA accessible to RNA signals, facilitating *de novo* methylation by DNA methyltransferases or demethylation (Kanno *et al.*, 2004; Kanno *et al.*, 2005).

The siRNA heterochromatic pathway in *Arabidopsis* is hypothesized to be a circular nuclear process that initiates with pol IVa activity at target DNA sites, which are typically localized in the chromocenter peripheries (Onodera *et al.*, 2005; Pontes *et al.*, 2006). RNA generated by pol IVa moves to a siRNA processing center in the nucleolus (Pontes *et al.*, 2006), that is suspected to be a Cajal body-like structure (Li *et al.*, 2006), where it is converted to dsRNA by RDR2 and diced by DCL3 to generate 24nt siRNAs. Within the nucleolar processing center, siRNAs are loaded into an effector complex that includes AGO4 and NRPD1b which then leaves the nucleolus and guides the effector complex to target DNA *loci*. In a sequence of events still not well understood, NRPD2a associates with the effector complex to form functional pol IVb and DRD1 and DRM2 proteins are then needed, in addition to pol IVb, to impose heterochromatic marks to the target chromatin (Li *et al.*, 2006; Pontes *et al.*, 2006).

I.3. *Nucleolous organizer regions (NORs)*

Ribosomal rRNA gene *loci* (NORs) transcription is specifically dependent on the activity of RNA polymerase I (pol I) (Albert *et al.*, 1999; Grummt and Pikaard, 2003; Hannan *et al.*, 1998; Saez-Vasquez and Pikaard, 1997). The pol I transcription activity in the NORs results in the most prominent

feature of the interphase nucleus, the nucleolus, a compartmentalized domain with low chromatin density. The nucleolus is where assembly of ribosomal particles takes place from the four ribosomal RNAs (rRNAs) transcribed by pol I (18S, 5.8S and 25S rRNAs) and RNA polymerase III (5S rRNA), and approximately 85 proteins whose mRNAs are transcribed by RNA polymerase II (Scheer and Weisenberger, 1994; Shaw and Jordan, 1995). The ultrastructural organization of the nucleolus is characterized by three distinct regions visible by electron microscopy: the fibrillar center, the dense fibrillar component, and the granular component (Melese and Xue, 1995; Olson *et al.*, 2000; Scheer and Hock, 1999). In the morphologically distinct fibrillar center, the rRNA genes are transcribed exclusively by pol I (Carmo-Fonseca *et al.*, 2000; Paule and White, 2000; Shaw and Jordan, 1995).

I.3.1. ribosomal RNA genes

In all eukaryotic cells, rRNA genes are organized as head-to-tail clusters on one or several chromosome *loci*. NORs contain a subset of active rRNA genes, that give rise to a secondary constriction on metaphase chromosomes, and silent genes that are compacted in heterochromatin. An intergenic spacer separates the standard rRNA gene transcription unit (Brown and Dawid, 1969; Reeder, 1974), being oriented such that the direction of transcription is toward the centromere (Copenhaver and Pikaard, 1996a and 1996b; Lin *et al.*, 1999; Mayer *et al.*, 1999; The Arabidopsis Genome Initiative, 2000). The first primary transcript of rRNA genes is a large 45-48S rRNA precursor, which is cleaved to give rise to 18S, 5.8S and 25-28S functional rRNAs.

Unlike most other gene families that evolve independently, every rRNA gene remains virtually identical within a population (Pikaard, 2000; Pikaard and Lawrence 2002). However, even in closely related species rRNA gene sequences can vary substantially, especially the non-coding regions such as intergenic spacers. In eukaryotes, rRNA gene intergenic spacers are typically dominated by the presence of repeated elements. Functional elements include the gene promoter, transcription terminators, repetitive enhancer elements and duplications of the gene promoter that are known as "spacer promoters" (Reeder, 1989). rRNA gene promoter activity measured in *A. thaliana* protoplasts showed that sequences between –33 upstream and +6 downstream of the transcription start site (defined as +1) are sufficient to program accurate pol I transcription *in vivo* (Doelling *et al.*, 1993; Doelling and Pikaard, 1995). In plants, the rRNA gene start site consists of the highly conserved TATATA(A/G)GGG sequence motif. Mutations on this consensus region abolish or severely inhibit transcription initiation suggesting that the conserved core sequence plays a role in both start site selection and promoter strength (Doelling and Pikaard, 1995). Duplications of the gene promoter are found upstream in the intergenic spacer of *Xenopus laevis* (Moss and Birnstiel, 1979), *Drosophila*, *Arabidopsis*, and mouse, and program transcription initiation that in Xenopus and flies has a positive

effect on pol I transcriptional activity at the downstream gene promoter. However, in *A. thaliana*, no significant enhancer activity was observed as a result of these promoter duplications (Doelling *et al.*, 1993; Doelling and Pikaard, 1995).

The rRNA gene sequence uniformity within a species but sequence divergence across species boundaries is known as concerted evolution (Dover and Flavell, 1984; Flavell, 1986; Gerbi, 1986). Unequal crossing over and/or gene conversion events are thought to be the mechanisms responsible for rRNA gene homogenization and concerted evolution (Dover, 1982), an idea for which there is some supporting evidence, especially in yeast (Petes, 1980; Szostak and Wu, 1980). However, with the exception of closely related species, eukaryotic rRNA promoter sequences have diverged significantly. Consistent with this sequence disparity, rRNA transcription is generally specific to taxonomic orders, the promoter of one group not being recognized by the transcription machinery of the others (Grummt *et al.*, 1982; Voit and Grummt, 1995).

In the case of *A. thaliana*, the fine structure of the NORs was deduced by two-dimensional pulsed-field gel electrophoresis, adding important insights into how gene homogenization and concerted evolution must proceed. Based on relative locations, at least four classes of rRNA gene variants exist in the ecotype Landsberg (Copenhaver and Pikaard, 1996b), defined by differences in the lengths of their intergenic spacers and the presence of a polymorphic HindIII restriction endonuclease site. All the genes at NOR2 were found to have a single intergenic spacer length whereas NOR4 is composed of three spacer length variant classes. These different variants at NOR4 are not intermingled or randomly distributed throughout the NOR. Instead, variants are clustered, with long variants at one end of the NOR, short variants at the other and intermediate-length variants in the middle. These observations suggest that the homogenization processes responsible for concerted evolution tend to act at short-range, leading to local spreading of rRNA gene variants (Copenhaver and Pikaard, 1996b). However, occasional gene conversion and/or unequal crossing over events between NORs must occur to allow for the concerted evolution of all rRNA sequences within the genome.

I.3.2. rRNA transcription and chromatin structure

Depending on the cell's need to manufacture ribosomes and produce proteins, rRNA synthesis can account for as much as 40-50% of the total transcription activity in actively growing cells, (Jacob, 1995, Warner, 1999), whereas in non-growing cells rRNA gene transcription is extremely reduced and is barely detected. Changes in pol I transcriptional activity are believed to mediate regulation of ribosome particle production and thus the potential of a cell to proliferate. Several reports indicate that in growing cells one subset of rRNA genes displays regularly spaced nucleosomes, corresponding to inactive gene copies, while others are in a transcriptionally open chromatin conformation. Regulation

of pol I transcription occurs at these transcriptionally competent rRNA gene copies mediated by changes of transcription rate (Grummt and Pikaard, 2003). Studies performed in two yeast strains demonstrated that overall mRNA transcriptional level remained unchanged despite one strain possessing three times more copies of rRNA genes than the other. The latter compensated for the reduced number of rRNA genes with a higher occupancy of pol I polymerases *per* gene (French *et al.*, 2003). A dramatic example of rRNA dosage control was reported in chicken cells that although polysomic for NOR bearing chromosomes displayed the same level of rRNA synthesis (Muscarella *et al.*, 1985).

By *in vivo* psoralen crosslinking, transcriptionally active and inactive genes were distinguished in exponentially growing mammalian cells that synthesize high levels of pre-rRNA. Only half of the rRNA genes are transcriptionally active in an open chromatin conformation, while the other half corresponding to inactive gene copies resides in a compact nucleosomal structure. The percentage of open, psoralen-accessible genes is reduced as cells enter stationary phase (Dammann *et al.*, 1993; Sandmeier *et al.*, 2002). A basal transcription factor Rrn3, a yeast protein homologue of mammalian transcription initiator factor TIF-IA (Dammann *et al.*, 1993; Sandmeier *et al.*, 2002) has been indicated as a player in the regulation of rRNA transcription during stationary phase. Rrn3 is thought to have a central role in the growth-dependent regulation of rRNA gene transcription, directly binding to pol I and thereby rendering pol I competent for initiation (Grummt and Pikaard, 2003). The complex of Rrn3-pol I is required for at least partly bridging pol I to the core factor that is bound to the gene promoter (Peyroche *et al.*, 2000). Although it is not clear how yeast cells select the ribosomal genes to be active or inactive, occurrence of the open chromatin structure on the coding region requires actively transcribing pol I molecules (Dammann *et al.*, 1995). It was proposed that disruption of the chromatin structure at active genes is mediated by pol I molecules as they advanced through the rRNA gene template (Grummt and Pikaard, 2003). This was supported by studies which show that in silent rRNA genes repeats, the entire transcription unit is packaged in compact nucleosomal cores, whereas in transcribed genes, the chromatin structure is extended (Prior *et al.*, 1983). All four core histones remain associated with the transcribed region, but in contrast with those associated with intergenic spacers or silent rRNA gene copies, the nucleosomes are unfolded. It is not understood why the process of transcription leads to unfolding of nucleosomes on pol I transcribed genes, whereas passage of RNA polymerase II (pol II) only causes a transient loss of one H2A/H2B dimer (Kireeva *et al.*, 2002).

In mammalian cells, rRNA genes are extensively methylated (Santoro and Grummt, 2001; Stancheva *et al.*, 1997). Transcription initiation complex formation requires binding of UBF, which is impaired if a single DNA base position (-133) is methylated (Santoro and Grummt, 2001).

Heterochromatinization of excess rRNA genes copies is achieved by action of NoRC, a complex containing TIP5 and the ATPase SNF2 (Strohner *et al.*, 2001). NoRC, through the TIP5 bromodomain, binds H4^{acetyl}K16, recruits DNA methyltransferase DNMT1 and the histone deacetylase HDAC1 (Santoro *et al.*, 2002) to promote deacetylation of residues K5, K8 and K12 on histone H4 and DNA methylation (Zhou and Grummt, 2005). The latter is restricted to promoter regions which might be related to RNA mediated targeting. TIP5 was shown to bind RNA and to interact with a ~150nt RNA encoded in the intergenic spacer of rRNA genes. Failure to bind RNA is thought to impair TIP5 localization to rRNA gene promoters (Mayer *et al.*, 2005). Finally, NoRC function promotes alterations in nucleosomal positioning in the promoter region. In active rRNA genes, promoter bound nucleosomes cover position -157 to -2 (+1 as transcription start site). Action of NoRC apparently shifts the promoter associated nucleosome 25 nucleotides downstream, over the transcription start site, impairing transcription complex formation (Li *et al.*, 2006).

Arabidopsis displays two classes of rRNA genes, with one set having high levels of DNA methylation and the other being unmethylated and associated with pol I and constituting the expressed fraction (Lawrence *et al.*, 2004). This is a common trait of highly expressed genes in this organism, which include rRNA genes, and it is hypothesized that high occupancy of polymerases engaged in transcription *per* gene unit reduce chances for DNA methylation to occur (Zilberman *et al.*, 2007). Further chromatin marks associated with the active class of RNA genes include enrichment in H3^{trimethyl}K4 whereas silent genes display H3^{dimethyl}K9 and are compacted in a perinucleolar domain as part of the NOR bearing chromosome's chromocenter (Lawrence *et al.*, 2004).

I.3.3. Nucleolar dominance, an epigenetic phenomenon

The first description of the phenomenon that has come to be known as nucleolar dominance resulted from the observation in *Crepis sp.* hybrids that after crossing different species, the secondary constriction at NORs of one progenitor were lost (Navashin, 1934). Heitz showed that secondary constriction and satellite formation at metaphase was related to nucleolus formation during interphase and that nucleoli form at or very near the constrictions (Heitz, 1931). Subsequent convincing evidence was provided by McClintock, who first used the term 'nucleolar organizer' to describe those regions (McClintock, 1934). Nearly forty years later, NORs were shown to be sites where rRNA genes are clustered in hundreds or thousands of copies (Phillips *et al.*, 1971; Wallace and Birnstiel, 1966), with secondary constrictions representing the transcribed genes (Wallace and Langridge, 1971) and the adjacent heterochromatin composed of excess inactive rRNA genes. In fact, secondary constrictions are now known to correspond to the set of rRNA transcribed genes. The physical association of rRNA

genes with structural components of the nucleolus impedes chromosome condensation in that region (Wallace and Langridge, 1971).

Nucleolar dominance is a widespread phenomenon in nature, occurring in plants, insects and mammals (Pikaard, 2000). In short, when two species are crossed to form an interspecific hybrid, only the NORs of one progenitor will form nucleoli, resulting from the transcription of its ribosomal RNA genes set. This mitotically stable, but reversible state of the underdominant class is not caused by alterations in rRNA gene sequences and is thus a classic example of an epigenetic phenomenon. Studies on the ribosomal gene family have been conducted in polyploids at both the structure and expression levels. At the structural level NORs are a useful marker for systematic (White et al., 1990) as well as for genome evolution studies (Hamby and Zimmer, 1992). Although the biochemical and genetic mechanisms responsible for establishment and maintenance of nucleolar dominance are only beginning to be understood, there is no single hypothesis that can explain the phenomenon. Hypotheses include the idea that species-specific differences in rRNA gene sequences and/or pol I transcription factors lead to the preferential activation of only one set of rRNA genes (Grummt, 2003; Russel and Zomerdijk, 2005). Alternative hypotheses suggest that one set of rRNA genes is singled out for repression, involving chromosomal influences not specified simply by rRNA gene sequences (Chen and Pikaard, 1997).

The study of nucleolar dominance, especially in genetic hybrids provides further useful insights of how metabolically active cells control rRNA gene transcription in response to both general metabolism and specific environmental changes. Recent studies in *Brassica* (Chen and Pikaard, 1997a; 1997b) and in *Arabidopsis* (Chen *et al.*, 1998) provided an empirical demonstration of a causal relationship between allopolyploidy, epigenetic modification and changes in gene expression in plants. It was inferred that cytosine methylation and histone deacetylation are implicated as partners in the enforcement and selective silencing of rRNA genes in nucleolar dominance (Chen and Pikaard, 1997a; 1997b; Lawrence *et al.*, 2004). However, the critical targets of these chromatin modifications are unclear. Possibilities include the individual rRNA genes, the large chromosomal domains that encompass rRNA gene clusters (Lewis *et al.*, 2004), or other, possibly unlinked regulatory *loci* such as genes encoding *trans*-acting factors (Pikaard, 2000). Although it is not clear how sequences adjacent to a NOR can influence rRNA gene activity, the underdominant genes are selectively repressed by mechanisms acting at the chromosomal level. Underdominant rRNA genes can be transiently expressed in a hybrid cell upon transfecting a plasmid containing copies of the underdominant set, even though their chromosomal counterparts are repressed (Chen *et al.*, 1998), also pointing to mechanisms selectively acting on the chromosomally encoded genes. Furthermore, rRNA genes silenced in vegetative tissues are derepressed in reproductive organs, indicating not only a reversibility

of the phenomenon but also differential expression during development. However, the mechanism by which dominant and under-dominant rRNA genes are discriminated in newly formed hybrids, leading to the initial establishment of nucleolar dominance, is even less understood. Nevertheless, the involvement of chromatin covalent modifications, such as DNA methylation and histone post-translational modifications suggest that chromatin remodeling is involved in the re-establishment of repressive chromatin states following hybridization and allopolyploidy (Chen *et al.*, 1998; Chen and Pikaard, 1997a).

In *A. suecica* hybrids silent rDNA genes of *A. thaliana* display uniformly methylated promoters associated with H3$^{\text{dimethyl}}$K9 and appear as condensed heterochomatin in interphase nuclei. On the other hand *A. arenosa* NORs are organized in two ways; the active subset is unmethylated and associated with H3$^{\text{trimethyl}}$K4 and the silent class has methylated promoter regions and is associated with H3$^{\text{dimethyl}}$K9. Disruption of nucleolar dominance, by either the histone deacetylase inhibitor TSA or by induced DNA demethylation caused by incorporation of 5-Aza-dC, is accompanied by an association between a subset of *A. thaliana* rDNA promoters with H3$^{\text{trimethyl}}$K4, which correlates with transcriptional activation (Lawrence *et al.*, 2004). The role of histone deacetylation was further analysed by resorting to RNAi knockdown of potential histone deacetylases in *Arabidopsis*, including RPD3-like and SIR2-like HDACs as well as the plant specific HDT family. In the hybrid species context, RNAi is a powerful tool to analyze gene function as it allows targeting of specific genes and bypasses the problems of gene redundancy in polyploidy species (Lawrence *et al.*, 2003). Both HDT1 and HDA6 histone deacetylases were found to be involved in the establishment of nucleolar dominance. Knock-down of both genes caused reactivation of silent *A. thaliana* rRNA genes by reducing DNA methylation levels at promoter sequences, shifting their association from H3$^{\text{dimetyl}}$K9 to an association with H3$^{\text{trimethyl}}$K4 and acetylated histones. Reactivation of *A. thaliana* NORs is accompanied by decondensation of the heterochromatic region of the NOR in interphase nuclei, further linking this cytological behavior with rDNA gene expression (Earley *et al.*, 2006; Lawrence *et al.*, 2004; Pontes *et al.*, 2004). Despite our increasing knowledge of the nucleolar dominance phenomenon in *A. suecica,* the choice mechanism that marks *A. thaliana* NORs for targeted silencing is still elusive.

I.4. *Arabidopsis as model system*

I.4.1. *Arabidopsis thaliana*

Arabidopsis thaliana (L.) Heynhold (2*n*=2x=10, where *n*=the number of chromosomes in gametes and x=the fundamental chromosome number) is presently the leading plant system for genetic and

molecular studies. Major advantages include its small plant size and short generation time and its propagation by self-fertilization. Complete genomic DNA sequence is available (The-Arabidopsis-Genome-Initiative, 2000) as well as extensive publicly available data on its transcriptome and methylome (Yamada *et al.*. 2003; Lu *et al.* 2006; Tran *et al.*, 2005; Wang *et al.*, 2005; Zhang *et al.*, 2006; Zilberman *et al.*, 2007). More recently, great effort has been made by a number of research groups in characterizing RNAi metabolism and databases for *Arabidopsis* smRNA populations are available (Gustafson *et al.*, 2005; Lu *et al.*, 2006). Furthermore, extensive projects generating loss of function mutants through mutagenesis or transgenesis have knocked out the function of the majority of f A. thaliana genes, providing a valuable resource for conducting studies on gene function (Alonso *et al.*, 2003).

Until few years ago, *Arabidopsis* received very little attention for cytogenetic research in spite of its simple five chromosome karyotype, as described by Laybach (1907). The advent of molecular cytogenetics and the use of DNA-specific fluorochromes and specific DNA probes for direct detection of gene sequences on chromosomal targets make this system suitable for studies of chromosomal organization in interphase nuclei (Heslop-Harrison, 1998; Heslop-Harrison *et al.*, 2003; Fransz *et al.*, 2002). FISH analysis of *Arabidopsis* chromosome domains have demonstrated that individual interphase chromosomes are organized as heterochromatic chromocenters with emanating euchromatic loops (Fransz *et al.*, 2002). The chromocenters contain all major tandem repeats, including rRNA genes, and the majority of the dispersed pericentric repeats, including transposon and non-transposon repeats. In contrast, euchromatin is gene-rich and forms 0.2 – 2 Mbp sized loops that emanate from the chromocenter. The territory of an *Arabidopsis* chromosome, comprising on average 25 Mb of DNA with 5200 genes (The Arabidopsis Genome Initiative, 2000), contrasts with human chromosomes, which are larger but have a lower gene density due to higher heterochromatic content (Fransz *et al.*, 2003).

I.4.2. *Arabidopsis suecica* as a model allopolyploid

The emergence of *A. thaliana* as a model species for plant genetics has created interest in *Arabidopsis suecica* as a model allopolyploid, a selfing species native to northern Europe. The parental origin of the *A.suecica* chromosomes was demonstrated by DNA sequence analyses (Kamm *et al.*, 1995; O'Kane *et al.*, 1996; Price *et al.*, 1994) and by *in situ* hybridization to the repeat sequences of either *A. thaliana* or *A. arenosa* (also known as *Cardaminopsis arenosa*). The 26 chromosomes of allotetraploid *A. suecica* include a diploid *A. thaliana* genome chromosome complement (10 chromosomes) and a 2x chromosome complement (16 chromosomes) from the closely related autotetraploid, *A. arenosa* (2n=4x=32). Although very closely related in some aspects,

these two *taxa* nonetheless exhibit 5 to 8% divergence of nucleotide sequence in protein-coding genes (Hanfstingl *et al.*, 1994; Henikoff and Comai, 1998a) and 30 to 40% divergence in the 180-bp centromeric repeats (Martinez-Zapater *et al.*, 1986; Round *et al.*, 1997; Vongs *et al.*, 1993).

In *A. suecica* nucleolar dominance results in the silencing of the *A. thaliana* rRNA gene subset (Chen et al., 1998; Lewis and Pikaard, 2001). Interestingly, nucleolar dominance is not independent of ploidy or gene dosage. Though the normal 2:2 dosage of progenitor genomes is correlated with the silencing of the *A. thaliana* rRNA genes (Chen et al., 1998) backcrossing *A. suecica* to tetraploid *A. thaliana* yields progeny which have a 3:1 dosage of *A. thaliana*: *A. arenosa* genomes and show a reversal in the direction of nucleolar dominance such that *A. arenosa* rRNA genes become underdominant (Chen *et al.*, 1998). The reversal in the direction of nucleolar dominance and silencing of *A. arenosa* genes argues against the hypothesis that A. arenosa rRNA genes having a higher binding affinity for one or more transcription factors and argues instead for alternative mechanisms that dictate which NORs can be activated. In *A. suecica* the three protein coding genes adjacent to NOR4, the nearest of which is only 3.1 Kb away, remain active in hybrids (Lewis and Pikaard, 2001). This indicates that the mechanisms responsible for nucleolar dominance are restricted to the NORs and do not act on larger segments of NOR-bearing chromosomes. Moreover, genetic analyses point to NORs as the units of regulation in nucleolar dominance (Lewis et al, 2004).

I.5. Aims of this work

Our understanding of nucleolar dominance as an epigenetic phenomenon has been advanced by numerous studies published in recent years. Nevertheless, little is known of the choice mechanism responsible for the determination of which parental set of rRNA genes to silence. The endogenous RNAi machinery provides the basis for an attractive hypothesis to explore due to the potential for targeting specificity and the involvement of siRNAs in the establishment of epigenetic marks. The allopolyploid *Arabidopsis suecica* is the ideal model system for this study as it benefits from the available resources available for *A. thaliana* as well as transformation methods to induce knock-down of candidate genes by RNAi (Waterhouse et al. 2003; Lawrence and Pikaard, 2003).

The major aim of my study has been to test the hypothesis that smRNAs are required for the establishment of nucleolar dominance in *A. suecica*. To accomplish the proposed task my study focused on:

1. Identifying rRNA gene-homologous smRNAs, the genes responsible for their biogenesis and the RNAi-related pathway involved.

2. Using RNAi mediated knock-down to evaluate the function of genes involved in biogenesis of smRNAs corresponding to rRNA genes and their possible modulation of nucleolar dominance in *A. suecica*.

II – Materials and Methods.

II.1. *Plant material and growth conditions.*

II.1.1. *A. thaliana* lines

A. thaliana ecotypes (WT):

- Columbia (Col-0); Arabidopsis Biological Resource Center (ABRC) stock center.

- Landsberg erecta (Ler) ; ABRC stock center.

- Nossen (No-0); ABRC stock center.

- Wassilewskija (WS); ABRC stock center.

A. thaliana mutant lines, in [ ] genomic background:

1) Argonaute protein family members:

- *ago4-1* [Ler]; At2g27040; made available by Steve Jacobsen (Zilberman *et al.*, 2003).

2) DICER-like enzymes:

- *dcl1-7* [Ler]; At1g01040; made available by James Carrington (Golden *et al.*, 2002).

- *dcl2-1* [Col-0]; At3g03300; made available by James Carrington (Xie *et al.*, 2004).

- *dcl3-1* [Col-0]; At3g43920; made available by James Carrington (Xie *et al.*, 2004).

- *dcl4-1* [Col-0]; At5g20320; made available by Hervé Vaucheret (Gasciolli *et al.*, 2005).

3) Chromatin remodellers:

- *ddm1-2* [Col-0]; At5g66750; made available by Eric Richards (Jeddeloh *et al.*, 1999).

- *drd1-6* [Col-0]; At2g16390; made available by Marjorie Matzke (Kanno *et al.*, 2004).

4) DNA methyltransferases:

- *drm1/drm2* [WS]; made available by Steve Jacobsen (Cao *et al.*, 2003).

- *met1-1* [Col-0]; At5g49160; made available by Eric Richards (Vongs *et al.*, 1993).

5) RNA polymerase IV subunits:

- *nrpd1a*-3,8 [Col-0]; At1g63020; SALK_128428, _083051 – ABRC stock center (Onodera *et al.*, 2005).

- *nrpd1b-11* [Col-0]; At2g40030 ; SALK_029919 – ABRC stock center (Pontes *et al.*, 2006).

- *nrpd2a-2/nrpd2b-1* [Col-0]; At3g23780 / At3g18090; SALK_046208 / SALK_008535 (Onodera *et al.*, 2005).

6) RNA dependent RNA polymerases:

- *rdr1-1* [Col-0]; At1g14790; made available by James Carrington (Xie *et al.*, 2004).

- *rdr2-1* [Col-0]; At4g11130; made available by James Carrington (Xie *et al.*, 2004).

- *sgs2-1 (rdr6)* [Col-0]; At3g49500; made available by Hervé Vaucheret (Mourrin *et al.*, 2000).

7) Other:

- *axe1-5* (*hda6*) [DR5]; At5g63110; made available by Tom Guilfoyle (Murfett *et al.*, 2001).
- *hen1-1* [Ler]; At4g20910; made available by James Carrington (Chen *et al.*, 2002).
- *rnps1-1* [Col-0]; At1g16610; SALK_004132 - ABRC stock center.
- RNPS1-FLAG [Col-0]; 35S driven full cDNA clone of *RNPS1* with N-FLAG protein tag, available at Craig Pikaard's lab.
- *smd3-1* [Col-0]; At1g76300; SALK_025193 - ABRC stock center.
- *xrn4-5* [Col-0]; At1g54490; SAIL_847 - ABRC stock center (Souret *et al.*, 2004).

II.1.2. *A. suecica* lines

- LC1; made available by Luca Comai (Chen *et al.*, 1998).
- 9502; derived from accession 90-10-085-10, originating in Finland (Pontes *et al.*, 2003).
- AGO4-RNAi [LC1]; generated in this study.
- DCL3-RNAi [LC1]; generated in this study.
- DRM2-RNAi [LC1]; generated by S. Preuss, Craig Pikaard lab.
- HDA6-RNAi [LC1]; (Earley *et al.*, 2006).
- HDT1-RNAi [LC1]; (Lawrence *et al.*, 2004).
- HEN1-RNAi [LC1]; generated in this study.
- NRPD2-RNAi [LC1]; generated in this study.
- RDR2-RNAi [LC1]; generated in this study.

II.1.3. *A. arenosa* line

- 3651; natural accession, originating from Poland; made available by Steve O´Kane.

II.1.4. Germination, growth conditions and selection

Plants were grown on soil or germination media according to particular requirements of each line and downstream application of the plant material. Plants were either grown in the greenhouse or in growth chambers under controlled ambient conditions (25°C; 16 hours light; 8 hours dark). Germination was induced and synchronized at 4°C for 2-3 days in the absence of light.

Seed sterilization.

Seeds to be sterilized were placed in a sterile microfuge tube; typically 50μl of seed were used. 1 ml of 1/5 mixture of commercial bleach and absolute ethanol was added and tubes inverted occasionally during a period of 10 min. Sterilization solution was removed and 1 ml of absolute ethanol added to remove traces of bleach. The washing step was repeated and tubes were moved to a laminar flow hood. Absolute ethanol was removed and tubes were left open overnight in order to allow seeds to dry. Seeds could then be used immediately or stored at room temperature for several months.

Sterile culture of Arabidopsis plants.

The culture germination medium was prepared by weighing 4.4 mg/ml of Murashige-Skoog basal medium with Gamborg's vitamins (SIGMA) in sterile distilled water, pH was adjusted to 5.7-5.8 using a solution of 1M KOH and 3-4 mg/ml of agar (AgarGel™, SIGMA) added to the solution. The suspension was autoclaved (121°C; 20 minutes) cooled to ≈55°C in a waterbath and poured into Petri dishes in a laminar flow hood. For plants requiring selection, selective growth medium was obtained by adding antibiotic stock solution to the desired final concentration. The majority of transgenic plants used in this study were selected with kanamycin (50μg/ml). HDA6-RNAi transformants were selected with hygromycin (25μg/ml), HDT1- and DRM2-RNAi with Finale™ herbicide (5μl/ml). Antibiotic stock solutions (x1000) were stored at -20°C. After 2-3 weeks, where applicable, plants were transplanted to soil.

II.2. *Generation of A. suecica RNAi lines.*

II.2.1. Preparation of RNAi construct for transgenesis.

Synthesis of cDNA from *A. thaliana* Col-0 RNA extracts (500ng) was performed using SuperScript™ III reverse transcriptase (Invitrogen) with 250ng of random primers d(N)$_9$ (New England Biolabs), according to the manufacturer's instructions. 1μl of cDNA was used as template in the subsequent RT-PCR reaction. Amplification reactions were analyzed by gel electrophoresis (1% agarose; 50mM TAE – Sambrook and Russel, 2001). The mRNA regions for RNAi mediated knock-down of the targeted genes were selected after sequence alignment of its mRNA sequence to the sequences of other gene family members. Only mRNA regions displaying low homology scores were used for RNAi vector construction.

Amplicons were gel purified with GeneClean™ Turbo (Q-BIOgene) and cloned into pENTR-TOPO vectors (Invitrogen) according to the manufacturer's protocols. Following bacterial

transformation, transformant selection (kanamycin 50μg/ml) and plasmid DNA extraction (QIAprep® Miniprep – QIAGEN), plasmid clones were sequenced in order to confirm DNA insert identity. Clones showing 100% homology to the selected mRNA region of each gene to be targeted (The Arabidopsis Genome Initiative, 2001) were selected for subsequent cloning steps.

Table 2.1 – Gene specific primers used for amplification of cDNA of selected genes for the generation of RNAi constructs. To all forward primers CACC was added to the 5'end in order to allow directional cloning into de pENTR-TOPO vector.

	Forward (5'→3')	Reverse (5'→3')	Amplicon length
NRPD2a	GTCTGGCTCCTCTTCTTTGCGCTA	AATGCTCAATAGGCTTCAAATCACCAT	490bp
RDR2	CTCAATGCGCTTGTTCATGC	AAATCCGAGACATGCTCTGC	349bp
DCL3	GCCACCTTTCAGGCTTAT	CGGATGAGGTATTGCACTGA	490bp
HEN1	GCTGAAGCTTTGATTTTGGC	AACCTCTAAGCAAGTGCCGA	453bp
AGO4	ACTTGCTGAGAAGAAGGGGC	CCGCATAGCTGATCTCCACT	458bp

DNA inserts were recombined into pHELLSGATE 8 (Syngenta; Helliwell *et al.*, 2002; Wesley *et al.*, 2001) with LR clonase (Invitrogen) according to the manufacturer's instructions, transformed into DH5α competent cells and selected with spectinomycin (100μg/ml) in solid medium (Luria Bertani medium, according to Sambrook and Russel, 2001). Plasmid DNA was extracted from obtained bacterial colonies (QIAprep® Miniprep – QIAGEN) and digested with BamHI or XhoI in order to screen for correct orientation of the PDK intron sequence, which could become inverted as a result of insert exchange at the two recombination sites of pHellsgate 8 and no longer be recognized by the endogenous intron-splicing machinery (Helliwell *et al.*, 2002). Finally, selected clones were sequenced with pHELLSGATE 8 specific primers (27-5 (fwd): GGG ATG ACG CAC AAT CC; 27-3 (rev): GAG CTA CAC ATG CTC AGG) to reconfirm sequence identity and orientation of recombined inserts

II.2.2. *Agrobacterium* mediated *A. suecica* plant transformation.

<u>*Agrobacterium* transformation.</u>

Recombined pHELLSGATE 8 was transformed into GV3101 competent cells by electroporation using 2μl of plasmid prep / 50μl of GV3101 liquid stock, at 250μFD (capacitance); 400Ω (ohms); 1.5volt. All procedures were performed on ice or using ice chilled materials.

Immediately after electroporation, 1ml of LB medium was added to the cuvettes, transferred to a 15ml culture tube and incubated at 28°C; 200rpm for 2h30min. 80μl of the liquid culture were spread on YEP solid medium (50μg/ml gentamicin; 10μg/ml rifamicin; 100μg/ml spectinomycin) (Sambrook and Russel, 2001). Transformed colonies were allowed to grow for 2 days at 28°C.

<u>*A. suecica* LC1 transformation.</u>

LC1 plants were grown on soil (1 plant/pot) until reaching the 8-10 true leaf stage. Pots were transferred to a cold room (4°C + light) for a period of 2 weeks in order to induce and synchronize flowering. LC1 plants were subsequently transferred to the greenhouse. Once the stem of the first inflorescence reached 3-5cm, the inflorescence was excised in order to induce lateral shoots to emerge and maximize the number of inflorescences to transform, typically 2 weeks after excision. Ideally, in order to maximize plant transformation efficiency, plants to be transformed should present immature inflorescences and not many pollinated flowers or fertilized siliques.

Agrobacterium 2ml cultures (YEP medium; 50μg/ml gentamicin; 10μg/ml rifampicin; 100μg/ml spectinomycin), transformed with the RNAi vector, were grown overnight (28°C; 200rpm) and transferred to 2.5 liter flasks containing 400ml of the exact same liquid medium and selection conditions. These cultures were allowed to grow overnight until dense cloud-like swirls of cells become visible. Cultures were centrifuged for 10min, 2,500rpm at 4°C and resuspended in 300ml of 5% (w/v) sucrose solution. Immediately before dipping, Silwet™ L-77 detergent was added to a final concentration of 0.05% (v/v) to aid wetting of plant surfaces. Inflorescences were immersed in the *Agrobacterium* suspension with gentle agitation (Clough and Bent, 1998). Dipped plants (T_0) were transferred to the greenhouse in order to complete their life cycle and progeny (seed) was collected (T_1). Typically, 4-5 LC1 plants were dipped for each RNAi construct. Following sowing the T1 seeds, transformants were selected (50μg/ml kanamycin) as described in section II.1.4.

II.3. *Analysis of genomic DNA methylation levels.*

II.3.1. Genomic DNA extraction.

2x CTAB buffer:

- 2.8ml 5M NaCl
- 0.2g CTAB (hexadecyltrimethylammonium bromide)
- 1ml 1M TRIS-HCl, pH8.0
- distilled H$_2$0 to 10ml

before use, add β-Mercaptoethanol to a final concentration of 1% (v/v).

1) 1 gram of leaf tissue was ground with liquid nitrogen until a fine powder was obtained. Plant material was transferred to a 15ml FALCON™ centrifuge tube and 5ml of 2x CTAB buffer was added. Mixture was vortexed and incubated at 65°C for 30min.

2) Following incubation, two sequential chloroform:isoamyl alcohol (IAA) (24:1) extractions were performed, each followed by centrifugation at 8,000rpm; 10min.

3) gDNA was precipitated in a pre-chilled 15ml FALCON™ with 0.8vol isopropanol;0.1vol 3M CH$_3$COONa and mixing by gently inverting tubes several times.

4) Following a 30min incubation (-20°C) and 30min centrifugation (8,000rpm; 0°C), pellets were washed twice with 70% ethanol (r.t.) and resuspended in 5ml of 1xTE buffer (Sambrook and Russel, 2001) containing RNAse A (10μg/ml).

5) After 1hour at 37°C, in order to achieve complete RNA degradation, the sample was sequentially extracted with phenol:chloroform:IAA (25:24:1) and chloroform:IAA (24:1). gDNA was precipitated with 3vol 100% ethanol; 0.1vol 3M CH$_3$COONa (-20°C; 30min), washed with 70% ethanol (r.t.) and resuspended in 200μl of 1xTE buffer. gDNA samples were stored at 4°C.

II.3.2. Southern blot hybridization.

Capillary blot.

Genomic DNA (1μg) was digested overnight according to the specific requirements of the restriction endonucleases used. Samples were separated on a 1.5-2.0% EtBr (Ethidium Bromide) stained agarose gel (1xTAE). Once electrophoresis was complete, gels were depurinated for 10min by immersion in a 0.25M HCl solution, with gentle agitation. Gels were rinsed in dH$_2$O. In-gel denaturation of DNA was performed with 3M NaCl; 0.4M NaOH (2x30min). Transfer of DNA to

a positively charged Zeta-Probe nylon membrane (Bio-Rad) was performed by overnight alkaline downward transfer (8mM NaOH; 3M NaCl). Finally, the membrane was rinsed in 2xSSC; 0.1% SDS and UV crosslinked. Membranes were stored at -20°C until used.

DNA probe labeling.

A 5S rRNA complete repeat was used as probe. PCR was used to amplify the 5S rRNA insert from the pCT4.2 plasmid (Campell *et al.*, 1992). The amplicon was gel purified with GeneClean™ Turbo (Q-BIOgene) and stored at -20°C until used.

Probe labeling was performed according to the following protocol:

1) 2μl of 5S rRNA purified amplicon

    1μl of dN(6) random primers (1ng/μl) (New England Biolabs)

    dH$_2$O to 10μl

    Mixture was denatured (95°C) for 5min and transferred to a waterbath at 37°C.

2) Following the denaturation step. the following were added:

    3μl of 1mM dNTP mix (no dCTP)

    5μl of α-^{32}P dCTP (3,000Ci/mmol)

    2μl 10x Exo-Klenow buffer

    1μl of Exo-Klenow (5u/μl)

    were added to the mixture. Labeling reaction was performed at 37°C for at least 1 hour.

3) Unincorporated nucleotides were removed with Performa® DTR gel filtration cartridges (Edge Bio). Labeled probe was immediately used for Southern blot hybridization.

DNA:DNA hybridization

Nylon membranes were pre-hybridized at 60°C in 0.25M Sodium Phosphate Buffer (pH7.2); 1mM EDTA; 7% SDS (see Sambrook and Russel, 2001 for preparation of stock solutions) in a hybridization chamber for 1 hour. Labeled probe was then added directly to pre-hybridization mixture and hybridization performed overnight at 60°C.

Blots were sequentially washed (60°C) for 5min in 2xSSC, 2xSSC; 0.1% SDS (2x 10min) and finally with 0.1xSSC; 0.1% SDS for 10min. Blots were removed from hybridization bottles, rinsed in 0.1xSSC; 0.1% SDS and wrapped in Saran-Wrap®. Blots were exposed to BioMax MS film (Kodak) at -80°C.

II.3.3. AtSNI methylation assay.

AtSN1 methylation assays were performed on ~50ng of gDNA digested with HaeIII. Semi-quantitative PCR conditions were 2 min at 94°C; 94°C for 30s, 53°C for 30s, and 72°C for 30s (32 cycles); 72°C for 2min. with *AtSN1* primers: 5'-ACTTAATTAGCACTCAA ATTAAACAAAATAAGT-3' and 5'-TTTAAACATAAGAAGAAGTTCCTTTTTCATCTAC-3'. A region of the At2g19920 *locus* was used as control (a *locus* not cut by HaeIII) and amplified with the gene specific primers 5'-CACCCGAACAGTTGGAAGAAGAG-3' and 5'-GTGAGGAACCGG TCCATTATTGCT-3'. PCR reactions were resolved and visualized by EtBr staining following agarose gel electrophoresis (1.5% agarose; 1xTAE) (Onodera *et al.*, 2005).

II.4. Protocols for RNA analysis.

For information on how to handle and prepare materials and solutions for RNA work consult Sambrook and Russel, 2001.

II.4.1. RNA extraction.

RNA was extracted using the mirVana® miRNA Isolation kit (Ambion) according to the manufacturer's instructions. Variations of the protocol allow purification of total RNA or high molecular weight (HMW) (>500nt) and low molecular weight (LMW) (<500nt) RNAs. Typically, RNA extractions were performed with 250-300mg of plant tissue. According to the downstream application of the RNA, leaf or immature inflorescence tissue was used. Inflorescence tissue RNA was used only for *A. thaliana* smRNA analysis. RNA samples were stored at -80°C.

II.4.2. RT-PCR.

DNase treatment of RNA samples and cDNA synthesis.

RNA to be used for the generation of cDNA was pretreated with RQ1 DNase (Promega). 10μg of RNA was incubated with the DNase at 37°C for 1 hour, 2.5μl of RQ1 stop buffer (Promega) was added, followed by a 10min (65°C) incubation to inactivate the enzyme. The final reaction volume (50μl) was adapted from the manufacturer's instructions.

RNA was further purified by adding 10vol of Trizol (SIGMA). Samples were mixed by inverting tubes and incubated 5min at r.t. 0.2vol of chloroform was then added to the mixture, vortexed and centrifuged for 10min at 10,000rpm. The aqueous phase was transferred to a new 1.5ml centrifuge tube and precipitated with 1.25vol of 100% ethanol (30min on ice). Finally, RNA was recovered by a 30min; 13,000rpm spin, washed with 80% ethanol and resuspended in 20μl of RNAse free MilliQ

H$_2$O. RNA integrity was checked on an agarose gel (rRNAs should be readily visible and no smear should be apparent). The final concentration of RNA was measured in a ND-1000 spectrophotometer (NanoDrop).

Synthesis of cDNA was performed using SuperScript™ III reverse transcriptase (Invitrogen) with 250ng of random primers d(N)$_9$ (New England Biolabs) or custom made oligo d(T)$_{18}$ according to the manufacturer's protocol. 500ng of DNase treated RNA were used per reaction. No-RT controls were prepared alongside reactions to which reverse transcriptase was added.

<u>Identification of transcripts overlapping the 45S rRNA core promoter region.</u>

A set of primers was designed in order to amplify RNA species overlapping the 45S rRNA core promoter (table 2.2). Additionally, a set of primers was also designed to evaluate the transcription status of *locus* At3g43160, found to possess a predicted intron sequence with high homology to the 45S rRNA core promoter region (see also chapter III). RT-PCR reactions were performed with Platinum *Pfx* Taq polymerase (Invitrogen) according to the manufacturer's instructions. The reaction mix was supplemented with DMSO to a final concentration of 5%. PCR conditions used were: For the 45S rRNA promoter - 94°C 2min; 94°C 30s, 50°C 30s, 68°C 30s (35 cycles); 68°C 2min. For the At3g43160 *locus* - 94°C 2min; 94°C 30s, 58°C 30s, 68°C 30s (35 cycles); 68°C 2min. *Actin2* (At3g18780) mRNA was also amplified as a control.

Table 2.2 – Primers designed to screen RNAs which overlap the 45S rRNA core promoter region (see also figure III.5 – chapter III)

	forward (5'→3')	Reverse (5'→3')	Primer pair
45S rRNA core promoter region		GGCACCAACAGACATTGAAATGTC	A
	CAAATGACCTAGCTAGAGGTGTC	CACCCCCCTATATAGCTTAATAGC	B
		AAAGGGGGTTCCCCACGGACTG	E
At3g43160	CGAGGATTCATCGACCAGGAATT	AGTCAGTCGGACGGTCGGTTGGT	
ACT2	GTTCTCTCCTTGTACGCCAGTGG	GTGCAACGACCTTAATCTTCATG	

<u>Evaluation of 45S rRNA *loci* expression levels in *A. thaliana* RNAi pathway mutants.</u>

Total RNA extracted from the 2nd and 3rd pairs of true leaves of individual plants was used for cDNA synthesis. Two plants for each genotype were analyzed; individuals were at the 4th leaf stage when plant material was collected.

Quantitative real-time RT-PCR was performed with Taqman® probes (Applied Biosystems) in an Real-Time PCR system (Applied Biosystems model 7500). PCR reactions were multiplexed in order to amplify both 45S rRNA and Actin in the same reaction, the latter to serve as an endogenous control. cDNA input corresponds to a 10-fold dilution of a cDNA synthesis reaction with 500ng of RNA input.

Table 2.3 – Primers and probes used for quantitative analysis of 45S rRNA gene expression levels in RNAi pathway mutants.

Target	Fwd primer	Taqman probe	Rev primer
45S rRNA (ETSregion)	5'-TGACATGGATTCTTCGAGGCCT-3'		5'-CATGACACGCCCATTCTCTTCG-3'
		6FAM-TATATAACTTGTTCGCATGATATT-MGB-NFQ	
ACT2	5'-TCGGTGGTTCCATTCT-3'		5'-GCTTTTTAAGCCTTTGATCTTGAGAG-3'
		NED-AGCACATTCCAGATGTGGATCTCCAA-MGB-NFQ	

Reaction mix

cDNA	1.0μl
JumpStart™ Taq ReadyMix™ (SIGMA)	12.5μl
Reference dye (SIGMA)	0.25μl
45S rRNA (ETS) probe	0.0625μl
Actin probe	0.0625μl
ETS primers (10pmol/μl)	0.4μl
Actin primers (10pmol/μl)	0.4μl
MgCl$_2$ 50mM	1.0μl
dNTPs 10mM	0.5μl
MQ-H$_2$O	8.825μl ∑=25μl

PCR program

1st step:	50.0°C – 2min
2nd step:	95.0°C – 10min
3rd step:	95.0°C – 15sec
	60.0°C – 1min (x40)

Results were analyzed with 7500 System SDS software (Applied Biosystems).

Evaluation of target mRNA knock-down levels in *A.suecica* RNAi lines.

In order to evaluate knock-down efficiency of RNAi targeted genes in *A. suecica* LC1, primers flanking the RNAi targeted region of the mRNA were designed.

Semi-quantitative RT-PCR reactions were performed with Platinium Taq (Invitrogen) according to the manufacturer's instructions. PFK was used as an endogenous control (*6-phosphofructokinase*;

At4g04040). Typically, cDNA input for PFK RT-PCR reactions corresponded to a 10-fold dilution of the input used for target gene knock-down evaluation. RT-PCR reactions were resolved and analyzed by gel electrophoresis (1% agarose; 1xTAE; EtBr staining).

Table 2.4 – Primers used for semi-quantitative analysis of RNAi mediated knock-down efficiency in *A. suecica* LC1 transgenic lines.

	Forward (5'→3')	Reverse (5'→3')
NRPD2	GGGCTGAGAAGGTGTTTATAGC	GGTGCGGATATCTAGCATCTCC
RDR2	GATTGATGCATTTCTTCTCAAGCTG	GCATTTGCGGGAAGCTTGCTCC
DCL3	TCATCTTCATCGGCTGCAGGTTC	TGGTTCTGTCTTCACAACCATCTC
HEN1	AAGCAATGAAGAGATGGAGTCTG	CTCAACAGATGAAGCTTCCCGTC
AGO4	CGATAAACTCAAAGGAGATGGATTC	CAAGCAAGCAACCTTGTCTAGC
PFK	CGCCGGAATTTCGATCAATCCT	CGCCACGAAAACCAAACAGAC

II.4.3. S1 nuclease protection assay.

Probe labeling.

Plasmids pAt1 and Asup1.7 (Chen *et al.*, 1997a), containing both IGS and ETS regions of 45S rRNA genes of *A. thaliana* and *A. arenosa*, were digested with EcoRV and BspEII restriction endonucleases, respectively. 5'- restriction ends were dephosphorilated with Antarctic Phosphatase (New England Biolabs) according to the manufacturer's instructions. Plasmid DNA was precipitated with 1vol of CH_3COONa 3M; 2.5vol 100% ethanol and resuspended in water.

Approximately 3μg of linearized plasmid DNA were 5'-end labeled (γ-^{32}P ATP) with T4 polynucleotide kinase (New England Biolabs) according to the manufacturer's instructions. DNA was precipitated with 1/2vol of CH_3COONH_4 7.5M; 3vol 100% ethanol, pelleted by centrifugation, washed twice with 70% ethanol and resuspended in 30μl of SphI reaction mix (New England Biolabs). 10u of SphI endonuclease were added to the solution and the cleavage reaction performed for 3hours at 37°C.

The digestion reaction was resolved in a 5% polyacrylamide (Accugel 19:1 – GeneFlow Ltd.); 100mM TBE gel. Bands corresponding to the IGS:ETS regions of *A. thaliana* and *A. arenosa* 45S rRNA genes were identified by exposing the gel to X-ray film, excised with a razor blade and transferred to a 1.5ml centrifuge tube containing 1ml of 1xTE buffer. Elution of gel purified bands was

performed overnight at 37°C. γ-^{32}P ATP incorporation was measured in a scintillation counter and $1X10^6$cpm (counts per minute) was used per hybridization reaction.

S1 nuclease protection assay.

For the analysis of parental specific 45S rRNA gene transcription activity in *A. suecica* lines, 20μg of HMW (>500nt) RNA was precipitated together with *A. thaliana* and *A. arenosa* specific radioactively labeled probes (in separate tubes) by addition of 1/3vol CH_3COONH_4 7.5M and 2.5vol of 100% ethanol. Samples were centrifuged at 13,000rpm for 20min (4°C), pellets washed in 70% ethanol and resuspended in 30μl of S1 hybridization buffer (40mM PIPES pH6.4; 400mM NaCl; 1mM EDTA; 80% deionized formamide). Samples were incubated at 90°C for 15min and immediately transferred to a 37°C water-bath. RNA-DNA hybridization was allowed to occur overnight.

270μl of pre-warmed (37°C) S1 digestion buffer (5% glycerol; 1mM $ZnSO_4$; 30mM CH_3COONa; 50mM NaCl; pH4.5) containing 200u of S1 nuclease (Roche) was added to each reaction tube and incubation was performed for 45min. Reactions were stopped by adding 10μl of S1 stop buffer (10% SDS; 0.5mM EDTA) followed by a brief vortexing step. DNA/RNA hybrids were precipitated with 30μl of CH_3COONH_4 7.5M and 1ml of cold absolute ethanol (-20°C) followed by a 20min centrifugation (13,000rpm; 4°C). Pellets were washed in 70% ethanol and resuspended in 6μl of S1 loading buffer (90% deionized formamide; 10mM NaOH; 1mM EDTA; 0.1% (v/v) Bromophenol blue (1mg/ml); 0.1% (v/v) Xylene cyanol (1mg/ml)). Samples were boiled for 3min, chilled on ice and separated by polyacrylamide gel electrophoresis (10% polyacrilamyde (Accugel 29:1 – GeneFlow Ltd.); 100mM TBE; 7M Urea).

Finally, gels were vacuum dried for 2hours (80°C) and exposed overnight to a phosphor screen and analyzed using a Personal Molecular Imager FX™ system (BioRad).

II.4.4. RT-PCR CAPS (Cleaved amplified length polymorphism).

By taking advantage of a polymorphism in the ITS1 regions of *A. arenosa* and *A. thaliana* 45S rRNA genes, creating an extra Hha I restriction site in the *A. arenosa* genes, Lewis and Pikaard (2001) developed a CAPS based method to evaluate parental specific 45S rRNA transcription status in *A. suecica*.

cDNA, equivalent to 5ng of HMW RNA input, was amplified with Platinum Taq (Invitrogen) according to the manufacturer's instructions (Invitrogen). Semi-quantitative RT-PCR reactions were as follows: 95°C 2min; 95°C 30s, 60°C 30s, 68°C 45s (29 cycles); 68°C 10min, with forward primer 5'-GCGCTACACTGATGTATTCAACGAG-3' and reverse primer 5'-CGCACCTTGCGTTCA

AAGACTCGA-3' (25μl reactions). 5μl of the amplification reaction were separated on a 1xTAE agarose gel to verify amplification results.

20μl of 2x reaction mix with 10u of HhaI restriction endonuclease were added to the remainder of the PCR reaction and incubated at 37°C for 3hours. Finally, reactions were resolved in a 2% agarose; 0.5xTBE; EtBr gel.

II.4.5. Northern blot – smRNA analysis

Transfer of LMW RNA by semi-dry electroblotting.

Inflorescence or leaf LMW (<500nt) RNA samples (8-10μg) were mixed with 1vol of 2x Gel Loading Buffer II (Ambion), incubated 80°C for 5min, placed on ice and separated by polyacrylamide gel electrophoresis (20% polyacrylamide: bisacrylamide (Accugel 29:1 – GeneFlow Ltd.); 50mM TBE; 7M Urea). RNA was transferred to a MAGNACHARGE nylon membrane (0.22μm) (GE Osmonics) by semi-dry electroblotting (50mM TBE buffer) at 2.0mA/cm2 for 2 hours. Membranes were rinsed in 2xSSC; 0.1% SDS and UV crosslinked. Membranes were stored moist at -20°C until use.

RNA probe synthesis and labeling.

All probes used for smRNA northern hybridization were derived from custom made DNA oligonucleotides designed according to the instructions of the mirVana™ miRNA probe construction kit (Ambion). DNA oligonucleotides are designed to include in their 3'-ends the CCTGTCTC nucleotide motif which hybridizes with an oligonucleotide containing a T7 RNA polymerase promoter. Following incubation with DNA polymerase and dNTPs to fill in resulting 3' ends, a T7 driven transcription reaction generates a labeled RNA probe (α-^{32}P CTP) which is subsequently used for RNA:RNA hybridization. After generating the RNA probe, according to the manufacturer's instructions, unincorporated nucleotides were removed with Performa® DTR gel filtration cartridges (Edge Bio). Probe was stored at -20°C until use.

Table 2.5 – Oligonucleotides used for the generation of RNA probes for 45S rRNA core promoter region smRNA mapping. Where indicated, the position is numbered relative to the 45S rRNA transcription start site.

Probe	5'-Sequence – CCTGTCTC-3'
Atr(-150-100) sense	TACCAGAAAATAGGATTTAGTATCCTTATGATGCATGCCAAAAGAATTT
Atr(-150-100) antisense	AAATTCTTTTTGGCATGCATCATAAGGATACTAAATCCTATTTTCTGGTA
Atr(-100-50) sense	TCAAATTCCAAGTATTTCTTTTTTCTTGGCACCGGTGTCTCCTCAGACAT
Atr(-100-50) antisense	ATGTCTGAGGAGACACCGGTGCCAAGAAAAAAGAAATACTTGGAATTTGA
Atr(-50+1) sense	TTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAAGCTATATAGG
Atr(-50+1) antisense	CCTATATAGCTTAATAGCCCTTTTCCCTCTTGGCACCAACAGACATTGAA
Atr(+1+50) sense	GGGGTGGGTGTTGAGGGAGTCTGGGCAGTCCGTGGGGAACCCCCTTTTC
Atr(+1+50) antisense	GAAAAGGGGGTTCCCCACGGACTGCCCAGACTCCCTCAACACCCACCCC
Atr(+50+100) sense	GGTTCGGACTTGGGTAGCGATCGAGGGATGGTATCGGATATCGGCACGAG
Atr(+50+100) antisense	CTCGTGCCGATATCCGATACCATCCCTCGATCGCTACCCAAGTCCGAACC
Atr(+100+150) sense	GAATGACCGACCGTCCGGCCGCCGGGATTTTCGCCGGAAAACTTTTCCGG
Atr(+100+150) antisense	CCGGAAAAGTTTTCCGGCGAAAATCCCGGCGGCCGGACGGTCGGTCATTC

		position
A	GAAATGTCTGAGGAGACACCGGTG	-69 to -45
B	CAACAGACATTGAAATGTCTG	-55 to -35
siR435	CCTCTTGGCACCAACAGACATT	-44 to -23
C	AGCCCTTTTCCCTCTTGGCAC	-33 to -13
D	CCTATATAGCTTAATAGCCCTTTT	-22 to +3
siR1192	AACACCCACCCCCTATATAGCTT	-9 to +15
E	CCAGACTCCCTCAACACCCACCCC	+3 to +26
Atr(+40)	CGAACCGAAAAGGGGGTTCCCCACGGACT	+28 to +57

Table 2.6 – Other oligonucleotides used for the generation of RNA probes used in smRNA northern blot hybridization. Where indicated, position numbers are relative to the 45S rRNA transcription start site. (*) denotes probe will detect smRNA species in sense orientation.

Probe	Position	5'- Sequence – CCTGTCTC-3'
Atr(-10)*	-46 to -10	CAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATT
Atr(-10)	-46 to -10	AATAGCCCTTTTCCCTCTTGGCACCAACAGACATTG
Aar(-10)	-46 to -10	TTTGTCCATTTTTGGGTCTGGCACCAGTGGAGATGC
Atr(-10+30)	-14+28	CCCAGACTCCCTCAACACCCACCCCCCTATATAGCTTAATAG
25SrRNA	(25S 3'-end)	TTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAAGCTATATAGG
siR759	(45S rRNA)	GAAGTCTCGGACCTGGTCGACGAA
siR1003	(5S rRNA)	AGACCGTGAGGCCAAACTTGGCAT
siR255	(ta-siRNA)	TTCTAAGTCCAACATAGCGTA
AthCEN	(CEN 180bp)	TGTATGATTGAGTATAAGAACTTAAAACGCAACCGCATCTTAAAAGCCTAAGTAGTATTTC
AtSNI		CCTCTATCTGAGAGATTTACCACTGGGCCAACAACACGTTGG
AtCOPIA		TTATTGGAACCCGGTTAGGA
miR159		TTTGGATTGAAGGGAGCTCTA
miR163		TTGAAGAGGACTTGGAACTTCGAT
miR164		TGGAAAGTGACTACATCGGGG
miR171		TGATTGAGCCGCGCCAATATC
miR172		AGAATCTTGATGATGCTGCAT

Northern blot hybridization.

Membranes were pre-hybridized at 42°C in 50% deionized formamide; 0.25M Na_2HPO_4 (pH7.2); 0.25M NaCl; 7% SDS (Sambrook and Russel, 2001) in an hybridization chamber for 1 hour. Labeled probe was denatured (90°C, 5min) and added directly to the pre-hybridization mixture. Hybridization was performed overnight at 42°C.

Blots were sequentially washed (42°C) for 2min in 2xSSC, 2xSSC; 0.1% SDS (2x 10min), 0.5xSSC; 0.1% SDS (15min) and finally with 0.1xSSC; 1% SDS for 10min. Blots were removed from hybridization bottles, rinsed in 0.5xSSC; 0.1% SDS and wrapped in Saran-Wrap®. Blots were exposed to BioMax MS film (Kodak) at -80°C.

Membranes were stripped with 50% deionized formamide; 0.1xSSC; 0.1% SDS at 65°C for 2hours before further use.

II.5. *Cytological analysis.*

II.5.1. Leaf nucleus extraction

Nuclear extraction buffer (NEB)

       10 mM Tris-HCl   pH 9.5

       10 mM KCl

       0.5 M sucrose

       4 mM spermidine

       10 mM spermine

       0.1% Mercaptoethanol

1) Chop 1g fresh tissue in 1-2 ml of NEB until a fine suspension is achieved. Transfer suspension to a centrifuge tube and fix plant material for 20 min, on ice, by adding 1vol of formaldehyde (8% formaldehyde solution, pH7.0).
2) Filter the suspension sequentially through a 100-50-20μm mesh filter (Miracloth, Calbiochem).
3) Centrifuge 3 min 2.5rpm, 4ºC.
4) Resuspended pellet with 40μl of NEB.
5) Apply 3 μl onto a glass slide.
6) Check nuclei quality by adding a drop of DAPI (4'-6-Diamidino-2-phenylindole in CITIFLUOR antifade solution (1μg/μl)) and cover with a coverslip.
7) Slides can be stored at 4ºC for up to a month.

II.5.2. Immunostaining of proteins in plant nuclei

10xKPBS

    1.28M NaCl

    20mM KCl

    80mM Na_2HPO_4

    20mM KH_2PO_4 (pH 7.2)

Primary antibody incubation

1) Re-fix preparations in formaldehyde 30 min and wash 3X 5min in 1xKPBS; 0.1% Triton X-100.

2) Block preparation by placing 200 μl blocking solution (2% BSA; 1xKPBS; 0.1% Triton X-100) onto the slide and cover with a 22x22mm coverslip in a humid chamber for 30 min at 37ºC.

3) Wash blocking solution 3X 5min in 1xKPBS; 0.1% Triton X-100.

4) Add primary antibody solution (50-100μl per slide - typically 1:200 dilution) in blocking solution (2% BSA; 1xKPBS; 0.1% Triton X-100). For the RNPS1-FLAG protein mouse anti-flag (SIGMA) was used. H3^{dimethyl}K9 was detected with a rabbit polyclonal H3^{dimethyl}K9 antibody (AbCam). Cover with a 22x22mm coverslip and place in the humid chamber overnight at 4ºC.

Detection, Counterstaining and Mounting

1) Wash 3X 5 min in 1xKPBS; 0.1% Triton X-100.

2) Preparations were incubated with blocking solution for 30 min in a humid chamber 37ºC (2% BSA; 1xKPBS; 0.1% Triton X-100).

3) Anti-mouse TRITC (for detection of RNPS1-FLAG) or anti-rabbit FITC (for H3^{dimethyl}K9) (SIGMA) secondary antibody was diluted, typically 1:1000, in 1xKPBS; 0.1% Triton X-100, added to the slides, covered with 22x22 mm coverslip and incubate for 2h at 37ºC.

4) Wash 3X 10 min in KPBS and mount the slides in DAPI (1μg/μl in CITIFLUOR) and cover with a coverslip 22x40 mm. Keep at 4ºC.

II.5.3. DNA:DNA Fluorescent *in situ* hibridization (FISH)

Solutions and reagents

- Pepsin (10 mg/ ml in 10 mM HCl)

- RNAse A (100μg/ml in 10 mM Tris.HCl, pH8.0)

- 20X SSC (3M NaCl; 0.3M Sodium Citrate; pH7.5)

DNA probe labeling

A. thaliana specific 45S rRNA probe was obtained by labeling the plasmid pAt.2 (Pontes *et al.*, 2003) with biotin 16 - dUTP (Roche) by nick translation (Roche Kit) according to the manufacturer's instructions.

<u>Pretreatment of chromosome preparations</u>

1) Following immunolocalization, slides were re-fixed with 4% FAA (50% ethanol; 5% formaldehyde; 10% acetic acid).

2) Incubate slides for 10 min at 37°C, in a humid chamber, with 200 μl pepsin/each with a 22x22mm coverslip, wash 3X in 2xSSC for 5min.

3) Add 200μl of RNAse A to each slide, cover with a coverslip, incubate in a humid chamber for 1h at 37°C. Finally, wash 3X in 2XSSC for 5min.

4) Incubate 10min at room temperature in 4% formaldehyde. After this treatment wash the slides 3X in 2XSSC.

5) Dehydrate slides through a 70%, 90% and 100% ethanol series, 3 min each, and air-dry.

<u>*In situ* hybridization solution and hybridization conditions</u>

Table 2.7 - DNA:DNA *in situ* hybridization mixture.

Reagents	**Final concentration in mixture**
100% formamide	50%
20XSSC	2X
50% dextran sulfate (w/v)	8%
10% SDS (w/v)	0.1%
Salmon sperm (10μg/ml)	1.5μg/μl
DNA probe	200-250ng *per* slide
dH$_2$O	Adjusted to a final volume of 30μl

Add 30μl hybridization mixture to each slide and immediately cover preparation with a 22x22mm coverslip. Denature slides at 80°C for 5 min on a thermocycler (MJ Research PCT-100; BioRad). Slides were allowed to cool to 37°C with 5min incubations at 70°C, 60°C, 50°C. Transfer the slides to a humid chamber and incubate overnight at 37°C.

Post-hybridization washes and detection

1) Slides were rinsed in 2xSSC for 3min and subsequently with 20% (v/v) formamide; 0.1xSSC at 42°C for 10 min.

2) 2x 5min washes in 2xSSC at 42°C, and 2x 4xSSC; 0.2% Tween-20 at room temperature were performed.

3) Slides were blocked for 15 min at room temperature with 200μl/slide of blocking solution (5% BSA; 4X SSC; 0.2% Tween20).

4) Primary detection was performed with mouse anti-digoxigenin antibody (Roche) (1:250 in 5% BSA; 4X SSC; 0.2% Tween20) with the incubation performed in a humid chamber for 1h at 37°C.

5) Slides were washed with 2xSSC; 0.2% Tween 20, 2X for 3min.

6) Secondary detection was performed with rabbit anti-mouse antibody conjugated to Alexa 488 (1:400 in 5% BSA; 4X SSC; 0.2% Tween20) (Molecular Probes).

7) Slides were washed as in step 5.

8) Finally, one drop of DAPI (1μg/μl in CITIFLUOR antifade solution) was applied and preparations covered with a 22x40 mm coverslip. Slides were stored at 4°C.

II.5.4. RNA:RNA Fluorescent *in situ* hybridization (FISH)

RNA probe synthesis and labeling

RNA probes were labeled by *in vitro* T7 polymerase transcription (mirVana™ miRNA probe construction kit - Ambion) with digoxigenin-11-UTP RNA labeling mix (Roche) according to the manufacturer's instructions.

In situ hybridization solution and hybridization conditions

RNA:RNA hybridization was performed immediately after the last wash in the protein immunodetection protocol. Therefore slides had not been previously mounted with DAPI.

Table 2.8 - RNA:RNA *in situ* hybridization mixture.

Reagents	Final concentration in mixture
1M PIPES, pH8.0	0.1M
0.5M EDTA	10mM
50% dextran sulfate (w/v)	10%
5M NaCl	3M
Yeast tRNA (25μg/μl)	5μg
RNA probe	1μg *per* slide
dH$_2$O	Adjusted to a final volume of 50μl

1) After adding hybridization mix, slides were incubated overnight at 42°C in a humid chamber.

2) Slides were washed in 2xSSC for 5min (42°C) followed by a 10min wash with 1xSSC; 50% formamide at 50°C.

3) 2x 5min washes in 2xSSC at 42°C, and 2x 4xSSC; 0.2% Tween-20 at room temperature were performed.

4) Slides were blocked for 15 min at room temperature with 200μl of blocking solution (5% BSA; 4X SSC; 0.2% Tween20) per slide.

5) Primary detection was performed with streptavidin-Cy3 (SIGMA) (1:200 in 5% BSA; 4X SSC; 0.2% Tween20) for 1h at 37°C.

6) Slides were washed as in step 3.

7) Secondary detection was performed with biotinylated anti-streptavidin (Vector Laboratories) (1:200 in 5% BSA; 4X SSC; 0.2% Tween20) for 1h at 37°C.

8) Slides were washed as in step 3.

9) Tertiary detection was performed with streptavidin Alexa 546 (Molecular Probes) (1:400 in 5% BSA; 4X SSC; 0.2% Tween20) for 1h at 37°C.

10) Slides were washed as in step 3.

11) Finally, one drop of DAPI (1μg/μl in CITIFLUOR antifade solution) was applied and preparations covered with a 22x40 mm coverslip. Slides were stored at 4ºC.

II.5.5. Microscopy

Preparations were examined using a Nikon Eclipse E800i epifluorescence microscope, with images collected using a Photometrics Coolsnap ES Mono digital camera. Images were pseudocolored and processed with Adobe Photoshop software (Adobe systems).

II.6. *Characterization of rnsp1 and smd3 SALK lines.*

II.6.1. Genotyping

gDNA was extracted from leaf tissue and PCR amplified with Extract-N-Amp™ Plant PCR Kit (SIGMA) according to the manufacturer's instructions.

SALK_004132 (*rnps1-1*) was genotyped with primers fwd 5'-AAACCAAGTCGTGG CCGTCGTTC-3', or rev 5'-CGTTGAGGAGATGTCTCTCTTGG-3' in combination with the left border primer of the T-DNA insertion LBa1 5'- GCGTGGACCGCTTGCTGCAACT (SALK Institute).

SALK_025193 (*smd3-1*) was genotyped with fwd 5'-GCAAGGGAAAGAGCGCTTCACTAGG-3', rev 5'-TCTCTCTTCCCTCCATGTAGAG-3' in combination with the left border primer of the T-DNA insertion LBa1 5'- GCGTGGACCGCTTGCTGCAACT (SALK Institute).

II.6.2. Evaluation of mRNA levels

In order to evaluate if T-DNA insertion resulted in knock-out of the target gene, RT-PCR was used in order to determine mRNA levels in the homozygous (-/-) T-DNA lines.

For identification of *smD3* transcripts, primers fwd 5'-GCAAGGGAAAGAGCGCTTCACTAGG-3' and rev 5'-TCTCTCTTCCCTCCATGTCTTCGCG-3' were used. Target mRNA sequence is located downstream of T-DNA insertion.

For identification of *rnps1* transcripts primers, fwd 5'-AAACCAAGTCGTGGCCGTCGTTC-3' (A) and rev 5'-CGTTGAGGAGATGTCTCTCTTGG-3' were used to amplify target sequences upstream of the T-DNA insertion site. Primers fwd 5'-TATAGATCTCCTCCAAGGGGCTC-3' and rev 5'-CTCTCAGTGGCCTCTTAGGACTG-3' (B) were used to amplify RNPS1 mRNA downstream of the T-DNA insertion site. Finally, combination of primers A and B allowed analysis of mRNA species on either side of the T-DNA insertion site.

III – Identification and mapping of rRNA gene-derived smRNAs in Arabidopsis thaliana.

III.1. Database survey for smRNAs homologous to the rRNA gene sequence.

III.1.1. rDNA-homologous smRNAs in the ASRP database

In order to identify and characterize the distribution of siRNAs homologous to the complete DNA sequence of the rRNA gene unit we initially searched the ASRP database (*Arabidopsis* small RNA Project; www.asrp.cgrb.oregonstate.edu/db/; Gustafson *et al.*, 2005), a publicly available *Arabidopsis thaliana* smRNA database. A total of 8,954 smRNA hits to the rRNA gene query sequence were obtained, ranging from 21 to 25nt in size and having no more than 2 nucleotide mismatches or 2 G:U pairs, which allowed for the possibility of small sequence variation within potential rRNA gene *variants*.

The great majority of the retrieved rRNA gene smRNA hits (94.7%) are in the sense orientation and, except for the 25nt class, the different size classes are similarly represented. In contrast, antisense smRNAs are less frequent and each size class encompasses a higher degree of nucleotide mismatch when compared to sense smRNAs (Table 3.1).

Table 3.1 – Frequency of rRNA gene homologous smRNAs according to size class and orientation. In parentheses is the percentage of smRNAs containing mismatches relative to the total number of hits in each class (data source: ASRP database).

	21nt	22nt	23nt	24nt	25nt	Total (n)
Sense (%)	22.4 (0.0)	21.0 (0.8)	21.4 (1.3)	22.2 (2.4)	13.0 (3.4)	8486
Antisense (%)	41.2 (6.6)	22.4 (13.3)	12.4 (10.3)	22.7 (10.4)	1.3 (0.0)	468
Total (n)	2094	1885	1874	1989	1112	8954

The distribution of the smRNA hits to the rDNA sequence revealed that the great majority correspond to the 45S rRNA primary transcript (98.5%) region and are in the sense orientation; almost no antisense smRNA hits were identified in the various ASRP libraries. SmRNAs homologous to the intergenic spacer (IGS) region of the gene, which accounts for almost one third of the rRNA gene sequence, represent only 2.4% of the total (table 3.2).

When a characterization by size class and gene region is performed the predominance of sense *versus* antisense smRNAs in the 45S rRNA region is striking (fig. 3.1). It is also noteworthy that the number of identified hits for each of the 21 to 24nt size class sense smRNAs is very similar (table 3.3) and reflects the fact that in most cases for any given sequence position all smRNA size classes are represented. This observation is not true in the IGS region or with respect to antisense smRNA hits.

Most likely, the smRNAs matching the region that corresponds to the 45S pre-rRNA result from rRNA degradation, explaining their sense orientation.

Table 3.2 – Frequency of rRNA gene homologous smRNAs according to gene region and orientation (data source: ASRP database).

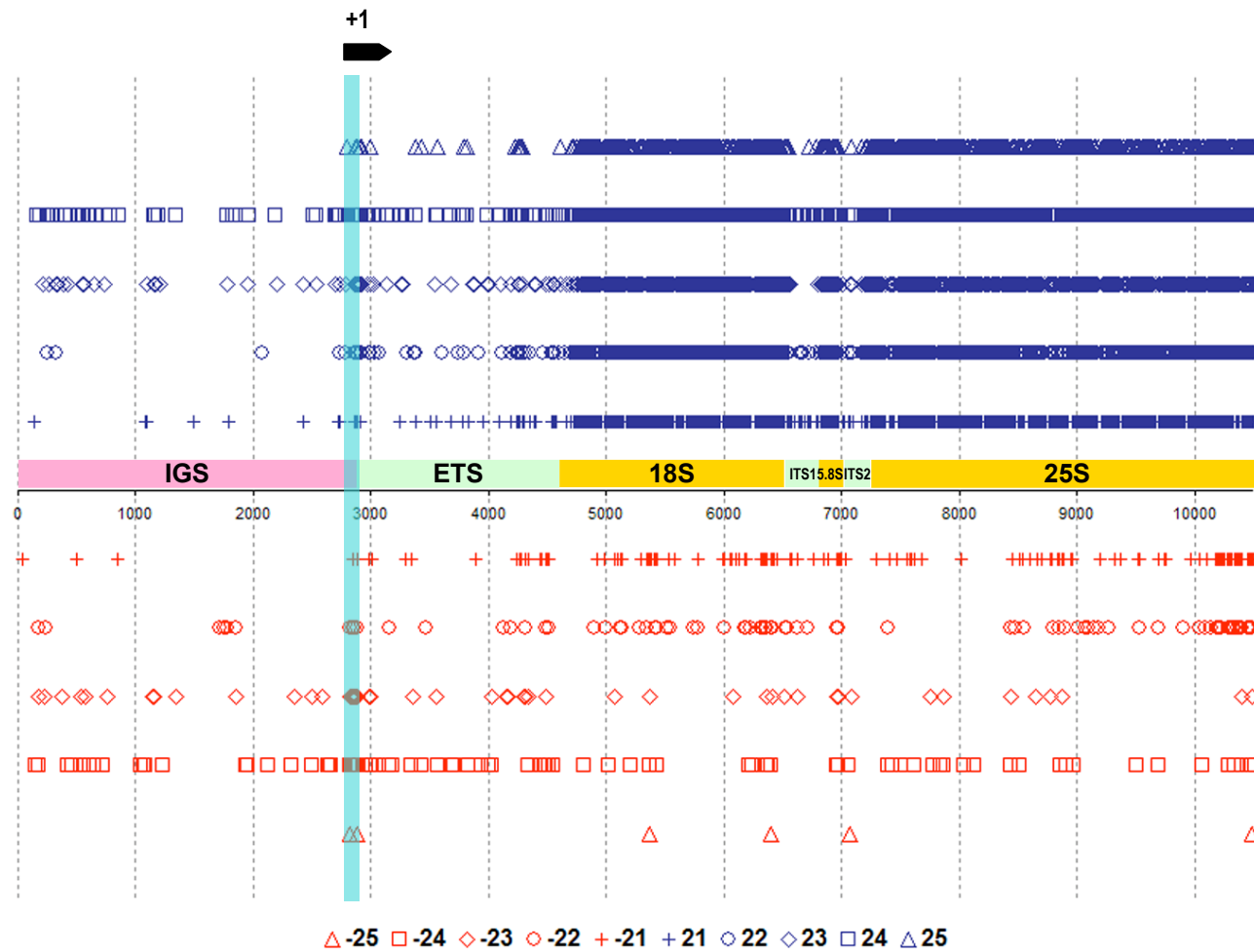|	IGS	ETS	18S	ITS1	5.8S	ITS2	25S	45S	Total (n)
Sense (%)	1.5	2.5	35.4	0.7	3.0	0.5	56.2	98.5	8486
Antisense (%)	17.9	13.5	23.1	1.7	3.2	3.2	37.4	82.1	468
Total (%)	2.4	3.0	35.1	2.4	1.6	0.8	54.7	97.6	8954

Table 3.3 – Distribution of smRNA hits by size class and rRNA gene region. (-) denotes antisense smRNAs. In parentheses is the ratio between sequence length / number of smRNA hits for a given region of the rRNA gene unit.

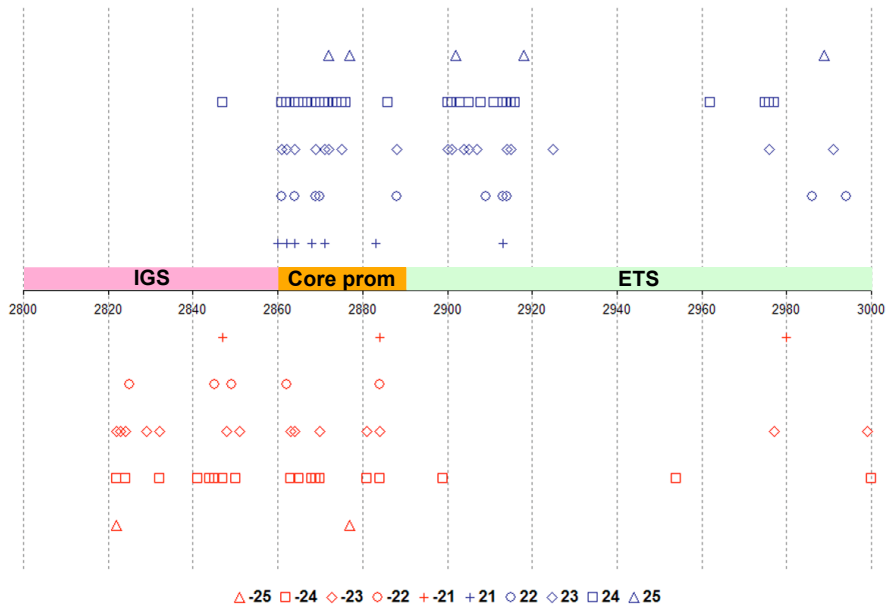| -25nt	-24nt	-23nt	-22nt	-21nt	21nt	22nt	23nt	24nt	25nt	rRNA region
2	40	25	12	5	14	10	34	65	4	**IGS**
(1446.5)	(72.3)	(115.7)	(241.1)	(578.6)	(206.6)	(289.3)	(85.0)	(44.5)	(723.0)	2893bp
0	26	11	8	18	42	37	44	63	23	**ETS**
(-)	(74.0)	(174.8)	(240.4)	(106.8)	(45.7)	(52.0)	(43.7)	(30.5)	(83.6)	1923bp
2	11	7	30	58	675	646	651	633	402	**18S**
(947.0)	(172.0)	(270.6)	(63.1)	(32.7)	(2.8)	(2.9)	(2.9)	(3.0)	(4.7)	1894bp
0	0	1	2	5	28	6	8	14	3	**ITS1**
(-)	(-)	(268.0)	(134.0)	(53.6)	(9.6)	(44.7)	(33.5)	(19.1)	(89.3)	268bp
0	4	2	3	6	57	51	59	61	39	**5.8S**
(-)	(41.0)	(82.0)	(54.7)	(27.3)	(2.9)	(3.2)	(2.8)	(2.7)	(4.2)	164bp
1	2	2	3	7	15	10	7	10	4	**ITS2**
(190.0)	(95.0)	(95.0)	(63.3)	(27.1)	(12.7)	(19.0)	(27.1)	(19.0)	(47.5)	190bp
1	23	10	47	94	1070	1020	1013	1037	631	**25S**
(3277.0)	(142.5)	(327.7)	(69.7)	(34.9)	(3.1)	(3.2)	(3.2)	(3.2)	(5.2)	3277bp
								n=8954		**(Σ)** 10525bp

Interestingly, the region that includes the core promoter of each rRNA gene (Doelling and Pikaard, 1993; 1995) is a location at which numerous smRNA hits occur, in both sense and antisense orientation, contrasting with the otherwise sparse distribution of smRNAs elsewhere in the IGS. Furthermore, by examining smRNA libraries originating from mutants defective for the four DICER-

Figure 3.1- Distribution of smRNA species according to size and orientation in a **A)** *A. thaliana* 45S rRNA gene repeat; **B)** 45S rRNA core promoter; **C)** 45S rRNA partial coding sequence. Data was retrieved from the ASRP database and queries were manually performed by searching for matches to a complete rRNA gene repeat. The search was restricted to 21 to 25nt smRNAs with no more than 2 nucleotide mismatches or 2 G:U RNA pairing hybrids. 3' ends were used to determine nucleotide positions of smRNA species for the purposes of the graphic display.
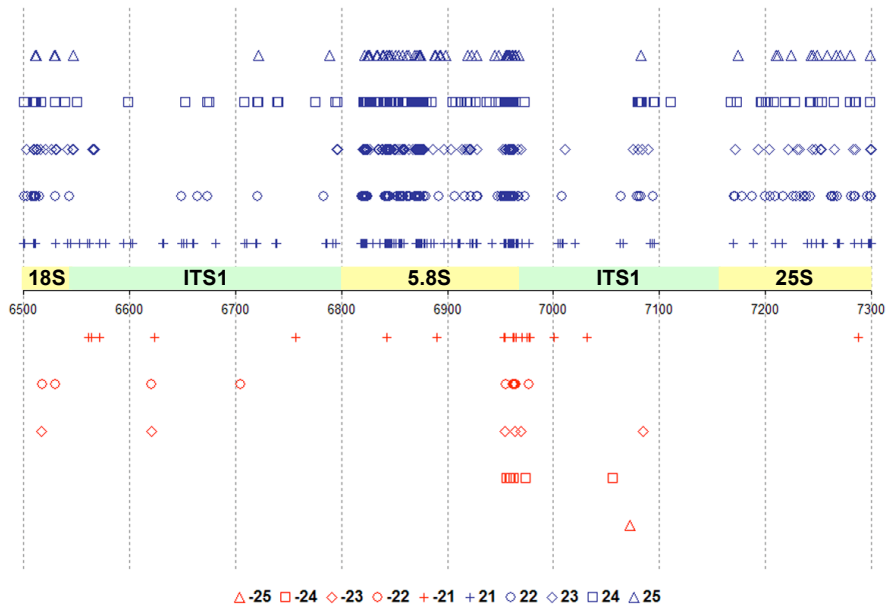
A)

△ -25   □ -24   ◇ -23   ○ -22   + -21   + 21   ◇ 22   ◇ 23   □ 24   △ 25

B)

IGS | Core prom | ETS

△ -25 □ -24 ◇ -23 ○ -22 + -21 + 21 ○ 22 ◇ 23 □ 24 △ 25

C)

18S | ITS1 | 5.8S | ITS1 | 25S

△ -25 □ -24 ◇ -23 ○ -22 + -21 + 21 ○ 22 ◇ 23 □ 24 △ 25

like enzymes or different RNA Dependent RNA polymerases, it is apparent that RDR2 and DCL3 are involved in the biogenesis of the rRNA promoter region 23-24nt class smRNAs (fig. 3.2). In *rdr2* mutants, almost no smRNAs corresponding to the IGS and ETS regions accumulate, revealing an important role for this RdRP in rDNA promoter smRNA biogenesis. This contrasts with the minimal involvement of RDR6. DCL1 and DCL4, both of which are responsible for the biogenesis of 21nt smRNAs, have little to no effect on promoter smRNA accumulation. The loss of DCL2 activity affects the 22nt size class but has no effect on 23-24nt promoter smRNAs. In contrast, smRNA accumulation in the 45S region is not visibly affected in any of the mutant lines, consistent with the hypothesis that these smRNAs are primarily due to rRNA turnover (not shown).

### III.1.2. Quantitative analysis of rDNA homologous smRNA

The limited number of smRNA hits to the rDNA gene repeat from the ASRP database provides relatively little quantitative information on the abundance of each smRNA identified. The use of Massively Parallel Signature Sequencing (MPSS) technology (Brenner *et al.*, 2000), and its adaptation to libraries enriched for smRNA (20-30nts) provides a better basis for estimating the abundance of different smRNAs, although in this case no size information is available as only 17nt *per* molecule are sequenced by this approach (Lu *et al.*, 2006).

Table 3.4 – Total number of smRNA hits by orientation and rRNA gene region. (-) denotes antisense; (+) sense orientation.

		IGS	ETS	18S	ITS1	5.8S	ITS2	25S	Total (n)
WT	+	394	75	315	28	30	7	580	1429
	-	393	77	63	20	12	3	272	840
rdr2	+	23	57	610	35	51	7	803	1586
	-	23	56	76	14	28	4	266	467

Access to WT (Col-0) and rdr2-1 mutant line MPSS data allowed the identification of distinct domains of RDR2-dependent smRNA accumulation in the IGS region (fig. 3.3a). A cluster of smRNAs matching the core promoter of rRNA genes was again observed, in agreement with the ASRP data described above. Additional regions of smRNA accumulation are found in the vicinity of the spacer promoters that are highly homologous to the gene promoter. Both sense and antisense smRNA species accumulate to similar levels and locate to the same regions of the rDNA IGS. The presence of siRNAs corresponding to both strands is consistent with the generation of siRNAs from dsRNA precursors, unlike smRNAs that might be derived from degradation of RNA coding

Figure 3.2- Distribution of smRNA species according to size and orientation in the *A. thaliana* rRNA gene core promoter region in smRNA libraries from mutant lines of known RNAi pathway genes. The graphic display shows 3kb, including the 3' end of the IGS and 5' end of ETS regions corresponding to an rRNA gene of *NOR2,* on the tip of chromosome 2. The transcription start site of the 45S rRNA is indicated (+1). Knockout of *RDR2* essentially eliminates all smRNAs matching the core promoter. *DCL3* is also involved in 23-24nt smRNA biogenesis. DCL2 is responsible for 22nt smRNA biogenesis. RDR6, DCL1 and DCL4 have no apparent role on production of core promoter smRNAs. No rRNA homologous miRNAs were found. Data were retrieved from the ASRP database (www.asrp.cgrb.oregonstate.edu/db/).
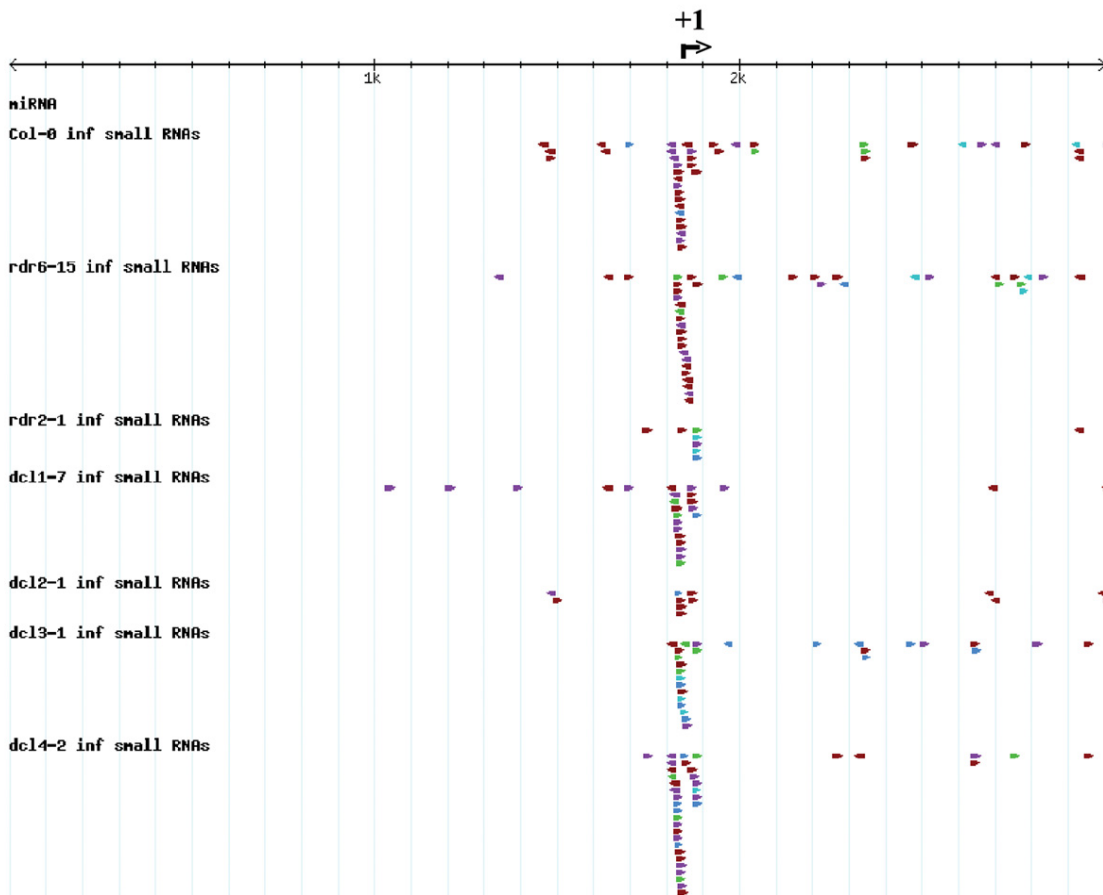
smRNA length color key: 19nt | 20nt | 21nt | 22nt | 23nt | 24nt | 25nt

sequences, which would be in the sense orientation only. Importantly, smRNA accumulation in the IGS is strongly reduced in the *rdr2* library demonstrating that rRNA IGS smRNA biogenesis is dependent on this RdRP. In contrast, RDR2 is not involved in smRNA biogenesis of the 45S region (fig. 3.3b; table 3.4). In general smRNAs exist throughout the entire coding sequence without any discrete preferential points of accumulation but with a clear trend towards higher smRNA accumulation levels in the sense orientation. It is interesting that the *rdr2* library actually shows an increase in sense smRNA accumulation especially in the regions corresponding to the functional rRNA subunits (table 3.4). Because RDR2 is the RdRP responsible for 24nt siRNA biogenesis one might speculate that the observed enrichment is a consequence of the reduction of RDR2 dependent smRNAs in the library such that the proportion of smRNAs generated by other pathways are over-represented.

## III.2. Validation of rDNA promoter region homologous smRNAs

The analysis of the distribution of smRNAs homologous to the rDNA genes revealed a cluster of smRNAs overlapping the rRNA gene promoter. In order to validate this observation northern blot hybridization was performed using low molecular weight RNA fractions (<500 nt) purified from wild-type *A. thaliana* ecotype Columbia. As a first step 50nt oligonucleotide probes were designed, in both sense and antisense orientation, homologous to 50 nt stretches of the rRNA gene sequence between positions -150 and +150 (with +1 defined as the transcription start site) (fig. 3.4a). The set of probes spanning -50 to +50 on both strands hybridized to 23-24nt smRNA species indicating that these are the predominant smRNAs size classes. Signal intensity was reduced between -100 and -50 and between +50 to +100 and was absent upstream of -100 or downstream of +100. Collectively, the data indicate that a 100bp region centered over the transcription start site, is a hotspot for smRNA accumulation.

To further characterize the smRNA population over the gene promoter, 22nt probes were designed, overlapping each other by 11nt, spanning positions -60 to +30. Two probes homologous to siR435 and siR1192, identified in the initial ASRP database screen, were used to set the phase for probe design. The 24nt siR1192, or siR1192-like smRNAs, was readily identified by northern blot hybridization (fig. 3.4b) but in contrast the siR435 probe failed to validate this 22nt siRNA; instead, the latter identified a 24nt smRNAs in agreement with the observation in figure 3.4a that 24nt siRNAs are the predominant size class. In the regions upstream the rRNA gene transcription start site, it is possible that the smRNAs are phased, such that not all 22nt probes are equally able to hybridize to smRNAs in this region whereas in the vicinity of the start site the smRNAs overlap each other such that all probes are equally effective. An additional probe covering positions +30 to +55 (Atr(+40)) did not hybridize

Figure 3.3- Accumulation levels and distribution of sense and antisense smRNA molecules in the *A. thaliana* rRNA gene repeat in WT and *rdr2-1* genetic backgrounds. **A)** The IGS region includes four locations where RDR2-dependent smRNAs accumulate in both the sense and antisense orientation. The three clusters located upstream of the core promoter (+1) correspond to a repetitive region located in the IGS where spacer promoters also localize (SP). The other discrete cluster overlaps the core promoter region and extends downstream into the ETS. **B)** Distribution and accumulation of smRNAs in the 45S rRNA coding sequence extending from the beginning of the 18S to the end of the 25S rRNA. Accumulation levels are higher in sense orientation and are evenly distributed. Nevertheless, regions near the 3' and 5' ends of 18S, 5.8S and 25S coding RNAs are regions for which corresponding smRNAs are particularly abundant. The data are derived from 17bp sequences of 20-30nt smRNA fractions obtained by Massively Parallel Signature Sequencing (MPSS) (http://mpss.udel.edu/at/; see also Acknowledgement section). The quantitative unit, TPQ (Transcripts per Quarter Million) is used to calculate the abundance of a given smRNA in each library in order to normalize results between different libraries (Gustafson *et al.*, 2005).
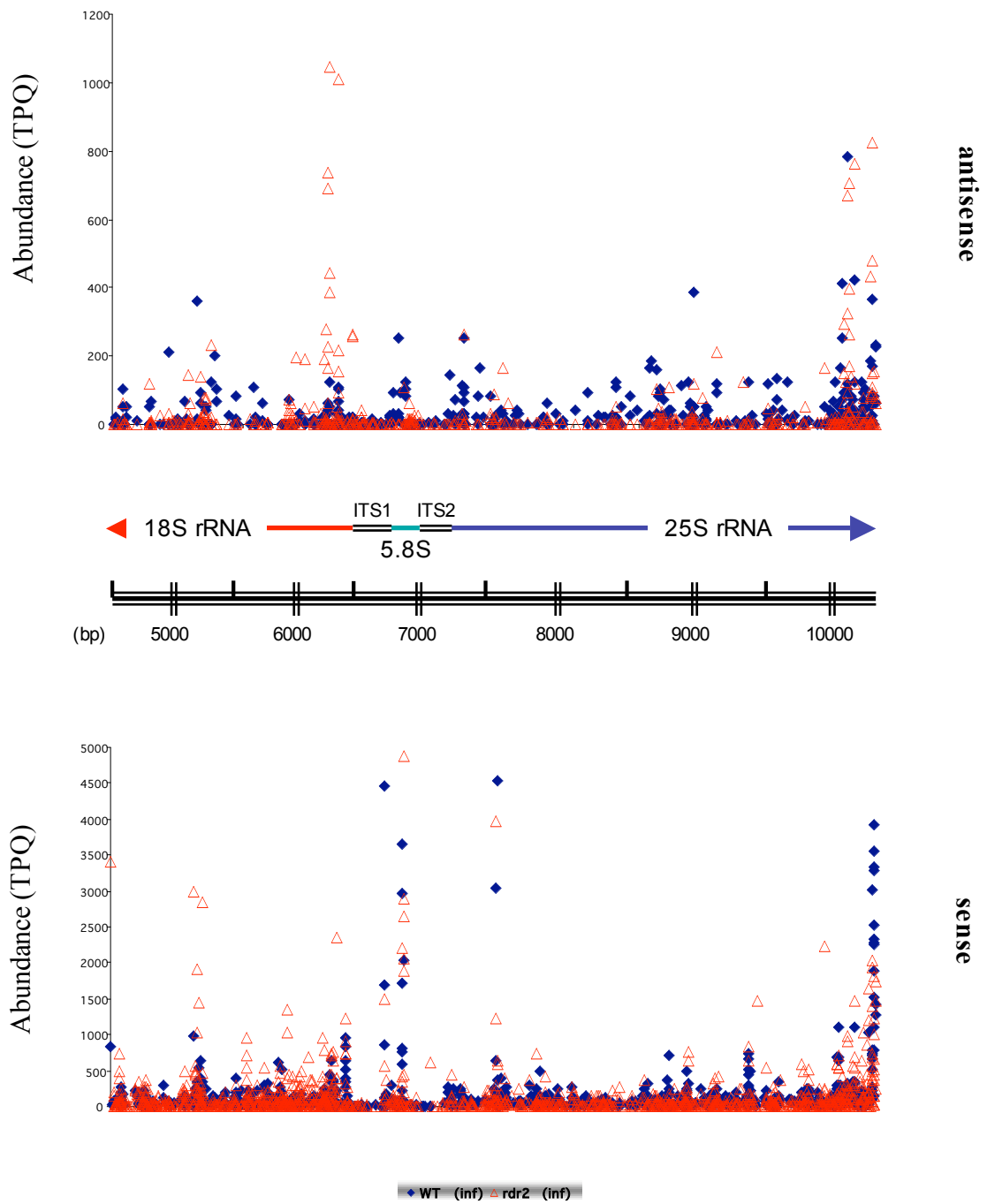
A)

rDNA smRNAs - IGS region
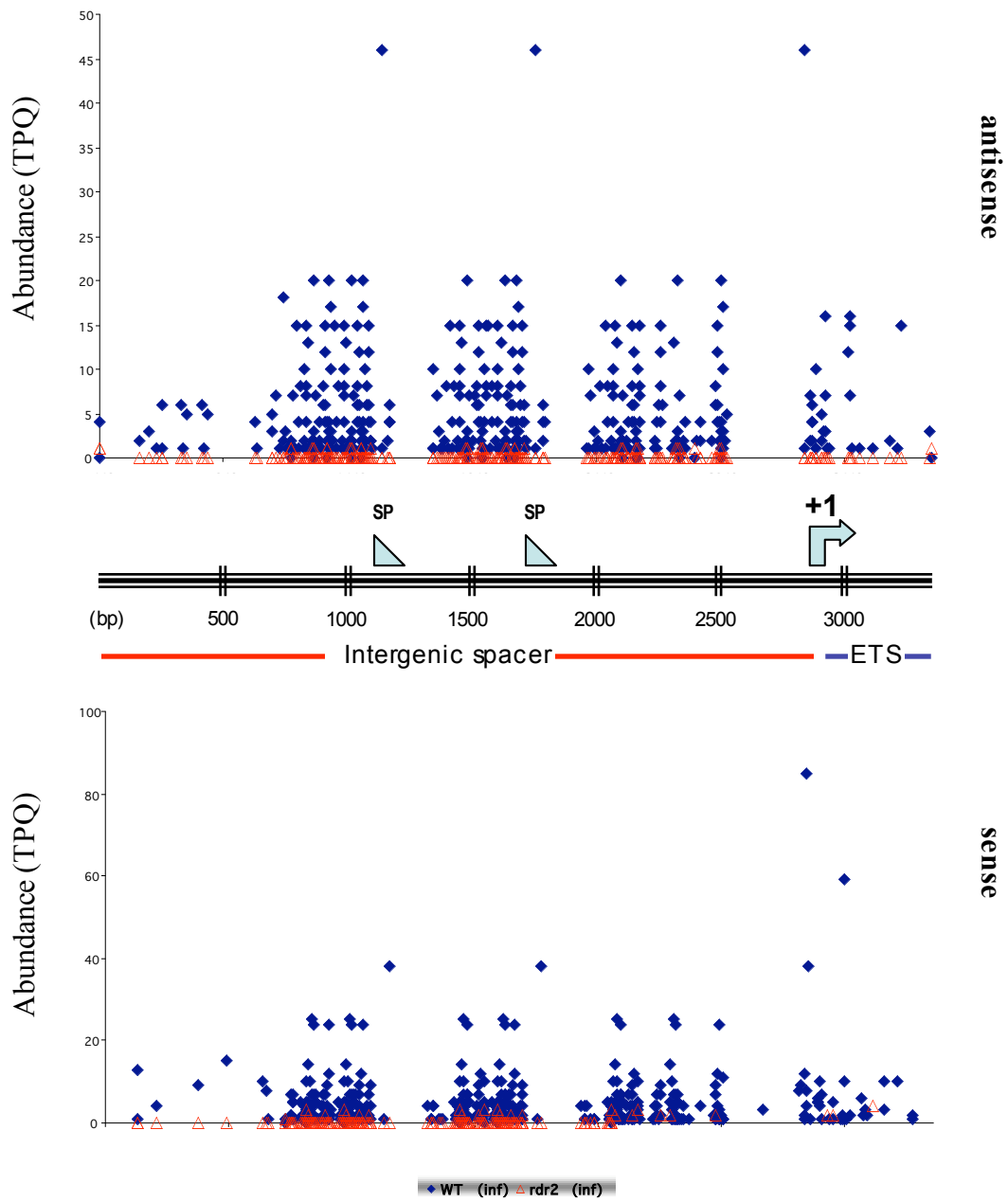
B)

rDNA smRNAs – 18S/25S region

strongly to 24nt smRNA species (fig. 3.4c). One can conclude that the interval between -60 and +30 is the preferential site for rRNA promoter smRNAs biogenesis.
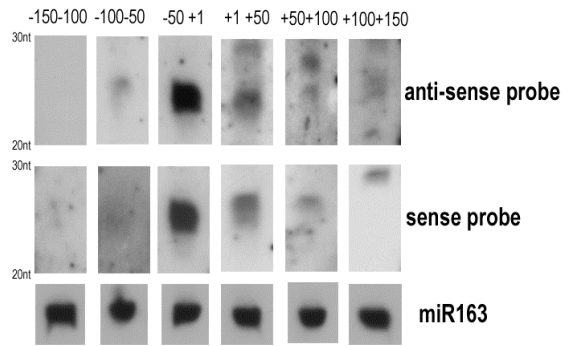
The distribution of smRNA over the rRNA gene promoter in smRNA libraries of different genetic mutant backgrounds, specifically mutants of the RdRPs and DICER-*like* enzymes known to function in *A. thaliana* RNAi pathways, indicated that RDR2 and DCL3 are required for the production of the abundant siRNAs corresponding to the rRNA gene promoter. As shown in Fig. 3.4 (panel C), DCL3 is the main DICER-*like* RNase responsible for rRNA gene promoter smRNA accumulation. As previously observed, the abundant 23-24nt smRNA corresponding to the gene promoter are essentially eliminated in the *dcl3-1* mutant line. In the *dcl3* mutant, low levels of 21-22nt smRNA molecules are observed and are likely due to the activity of other DICERs acting on the dsRNA precursors normally diced by DCL3. On the other hand, no significant alteration in smRNA accumulation was observed in *dcl2-1* relative to WT indicating no function for DCL2 in 24nt rRNA smRNA biogenesis. Likewise, *dcl1-7* and *dcl4-1* are not appreciably involved in the production of the 23-24nt size class rRNA siRNAs. Analysis of additional mutants validated the direct role of RDR2, and the non-involvement of RDR6, in rRNA promoter smRNA biogenesis (fig 3.4c). Of the analyzed RDR genes, only in *rdr2-1* was a loss of smRNAs observed in both sense and antisense orientation. Therefore, figure 3.4 shows that two principal RNAi pathway players involved in rRNA gene promoter siRNA biogenesis are RDR2 and DCL3.

SiR759, or related smRNAs with 1-2 mismatches relative to SiR759 sequence, are highly represented in vicinity of spacer promoter regions. Northern blot analyses similar to those shown in figure 3.4 showed that these IGS smRNAs also require RDR2 and DCL3 for their biogenesis.

Northern blot hybridization failed to validate smRNA species corresponding to 25S coding region antisense smRNAs (fig. 3.4c). One would expect that the high accumulation levels of coding region smRNAs relative to rRNA promoter siRNAs would allow the coding region smRNAs to be readily identified. Such was not the case, which may be due to the fact that coding region smRNAs are generated randomly, such that no one sequence is consistently generated. XRN4, the *Arabidopsis* homolog of *XRNP1p* of *S. cerevisae, which is known to be* involved in degradation of rRNA processing intermediates (Stevens *et al.*, 1991), seems to be involved in rRNA turnover in *Arabidopsis* as well because 21nt smRNA species corresponding to the 25S rRNA accumulated in *xrn4* mutants (fig. 3.4c). The same is not observed for IGS or ETS derived smRNAs, again suggesting that IGS and coding region smRNAs are generated in different ways.
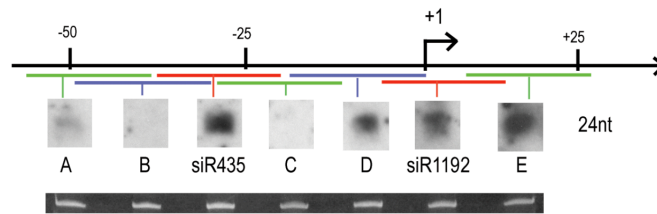
Figure 3.4- Validation of rRNA smRNAs by northern blot hybridization using radioactively labeled RNA probes. **A)** Tandem 50nt probes in sense and antisense orientation spanning the core promoter region from -150 to +150 (+1; transcription start site) validate the presence of smRNA species in this region. 23-24nt size class smRNAs are predominant, accumulate in both orientations and are most abundant in the -50 to +50 sequence interval. **B)** Probes against antisense smRNA species are 22-24nt in length and overlap each other by 11nt (for exact probe sequence coordinates see Chapter II). siR435 (22nt – see text in III.2.) and siR1192 (24nt), identified in the ASRP database, were used to set the phasing of the probes. Results indicate -60 to +30 as the major region of smRNA accumulation, with highest smRNA levels near the +1 transcription start site, which is overlapped by siR1192. **C)** RDR2 and DCL3 were validated as key players in rRNA homologous siRNA biogenesis. Both core promoter (Atr(-10) and Atr(-50+1)) and spacer promoter (siR759) siRNAs are essentially eliminated in *rdr2-1* and *dcl3-1* mutant lines. *RDR6* and the other dicers are not involved in their biogenesis. Failure to identify 24nt siRNAs with probe Atr(+40) (homologous to positions +27 to +55) suggests that most siRNAs accumulate upstream, nearer the rRNA gene core promoter as determined in **A)** and **B)**. 25S homologous siRNAs were not detected by RNA blotting except that the accumulation of 21nt smRNA species occurred in *xrn4*. MiR163 and siR255 are shown as loading controls.
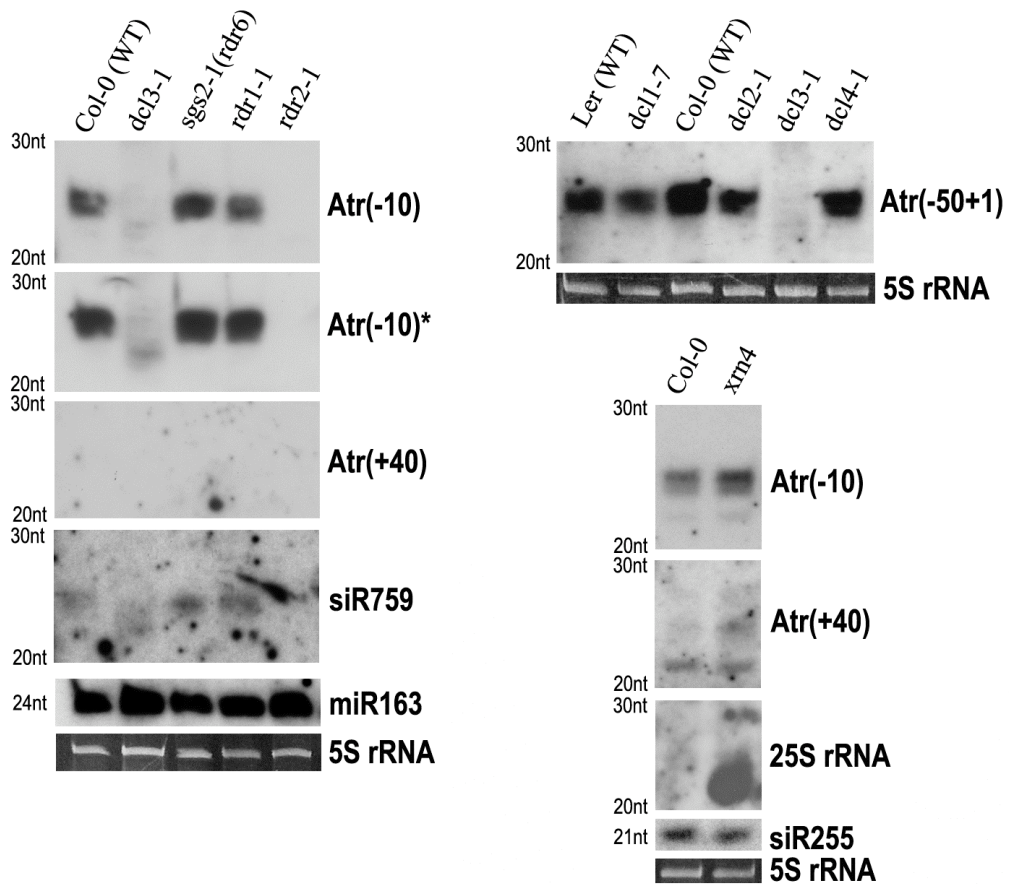
A)

-150-100   -100-50   -50 +1   +1 +50   +50+100   +100+150

anti-sense probe

sense probe

miR163

B)

rDNA gene promoter

-50   -25   +1   +25

24nt

A   B   siR435   C   D   siR1192   E

C)

Col-0 (WT)   dcl3-1   sgs2-1(rdr6)   rdr1-1   rdr2-1

Atr(-10)

Atr(-10)*

Atr(+40)

siR759

miR163

5S rRNA

Ler (WT)   dcl1-7   Col-0 (WT)   dcl2-1   dcl3-1   dcl4-1

Atr(-50+1)

5S rRNA

Col-0   xrn4

Atr(-10)

Atr(+40)

25S rRNA

siR255

5S rRNA

III.3. smRNA precursors originate from rRNA genes

The biogenesis of siRNAs implies the existence of a precursor RNA to serve as template for RNA dependent RNA polymerase activity and the dicing of the generated dsRNA molecule by DICER-*like* enzymes into 21 to 24nt short dsRNAs. The unwinding of this short dsRNA is the final step in the biogenesis of the functional single stranded smRNA molecule. If there are siRNAs overlapping the 45S rRNA gene promoter region it necessarily means that transcripts are being generated that encompass homologous upstream and downstream sequences of the transcription start site of rRNA genes.

Comparison of siR435 and siR1192 sequences against the Arabidopsis genome identified three other locations that could potentially be sources of RNA precursors for siRNAs homologous to the rDNA promoter region. BAC F21A14 on chromosome 3 contains an almost complete and highly homologous (97%) 45S rRNA gene repeat but with a truncated IGS restricted to the 550bp immediately upstream of the transcription start site. This truncated insertion is flanked by rearranged portions of IGS and 25S coding sequence. An intron of the At3g43160 *locus* is also highly homologous to the rRNA gene promoter and curiously has an annotated putative splicing site 17bp upstream of a homologous 45S rRNA transcription start site. Finally, BAC F14M2 (chromosome 1) consists of rearranged portions of IGS and ETS sequences (60-80% homologous to rDNA) but still conserves almost identical sequences to the rDNA promoter region. Therefore, besides NOR2 and NOR4, the NORs on chromosomes 2 and 4, respectively, other regions could potentially transcribe an rRNA precursor capable of triggering rRNA promoter derived siRNA biogenesis.

In order to identify RNAs overlapping the 45S rRNA transcription start site and their possible origin, primers flanking the 45s rRNA gene core promoter and the validated sites of siRNA accumulation were used in RT-PCR reactions (fig 3.5a). The amplicons obtained were gel purified and sequenced to verify which of the candidate *loci* could be responsible for transcribing the precursor RNA for rRNA siRNA biogenesis. The retrieved sequenced data showed near perfect homology to the rRNA sequences found in both NOR2 and NOR4 repeat clusters as well as the *locus* in BAC F21A14 (>98%), but not to BAC F14M2 or the At3g43160 *loci* (fig. 3.5c). For the latter, specific primers were designed in exons unrelated to rRNA sequences and flanking the intron whose sequence is homologous to the rRNA gene promoter. No RT-PCR product was obtained as a result of transcriptional activity of this gene leading to the conclusion that it is likely silenced (fig. 3.5b). As for the rDNA *locus* in BAC F14M2, the validation of 24nt siRNA similar to the annotated siR435 (fig. 3.4b), and the lack of homologous DNA sequence for this particular smRNA in the F14M2 sequence makes it unlikely that this locus is contributing for the 45S rRNA smRNA pool (fig. 3.5c). Although

Figure 3.5- Identification of RNAs overlapping the 45SrRNA transcription start site. Truncated NOR-homologous DNA sequences are present in other locations in the *Arabidopsis* genome and could be potential sources of   siRNA precursor RNAs **A)** RT-PCR of transcripts overlapping the transcription start site of 45S rDNA genes in *A. thaliana* ecotype Col-0; the primer positions are indicated in diagram. **B)** RT-PCR of At3g43160 failed to identify any transcripts (400bp) from this *locus* indicating that it is unlikely to be a source of precursor transcripts for 45S rDNA-homologous siRNAs shaded area in C corresponds to a putative exon). **C)** Sequence alignment of the RT-PCR amplicon obtained with the primer E and the different *loci* containing core promoter-related sequences. Highlighted is the position of siR435 in the core promoter sequence.

A)

400bp
200bp
100bp

A B E   A B E

+RT   -RT

E
B
A
+1
-186   -30 +7 +47

Primer annealing sites – 45S

B)

At3g43160   Actin

cDNA   d(N)$_9$ d(T)$_{18}$ d(N)$_9$ d(T)$_{18}$
RT   +   -   +   -   +   +   DNA

1000bp
500bp

At3g43160 specific primers

C)

```
RT-PCR     CAAATGACCTAGCTAGAGGTGTCAAAAAAATTATGAAAATTTACCAGAAAATAGGATTTAG 60
NOR2&4     CAAATGACCTAGCTAGAGGTGTCAAAAAAATTATGAAAATTTACCAGAAAATAGGATTTAG 60
F21A14     CAAATGACCTAGCTAGAGGTGTCAGAAAATTATGTAAATTTACCAGAAAATAGGATTTAG 60
At3g43160  CAAATGACCTAGCTAGAGGTGTCGGAAAATTACGTAAATTTACCAGAACATAGGATTCAC 60
F14M2      CAAATGACCTAGCTAGAGGTGTCAGAAAATAACGTAAATTTACTTGAAAATAGGATTCAC 60
           ***********************  ***** *  * ******** *** ******** *

RT-PCR     TATCCTTATGATGCATGCCAAAAAGAATTTTCAAATTCCAAGTATTTCTTTTTTTTTGGC 120
NOR2&4     TATCCTTATGATGCATGCCAAAAAGAATTTTCAAATTCCAAGTATTTCTTTTTTCTTGGC 120
F21A14     TATCCTTATGATGCATGCCAAAAAGAATTTTCAAATTCCAAGTATTTCTTTTTTTTTGGC 120
At3g43160  TATCCTTATGAAGCATGCCAAAAACAATTTTCAAATTCCAAGTATTTTTAATTTTTTGGC 120
F14M2      TATCCTTATGAAACTTGCCAAAAATAATTTTCAAATTCCAAGT---TTTTTTTTTTGGC 117
           ***********  * ********* ****************** *  * *** *****

RT-PCR     ACCGGTGTCTCCTCAGACATTTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAA 180
NOR2&4     ACCGGTGTCTCCTCAGACATTTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAA 180
F21A14     ACCGGTGTCTCCTCAGACATTTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAA 180
At3g43160  ACCGGTGTCACCTCAGAAATTTCAATGTTTGTTGGTGCCAAGTCGGAAAAAGGATATAAA 180
F14M2      ACCAGTAACTCCTTG-----------------TGGTGCCAAGTTGAAAAAAAGACATAAA 160
           *** ** *  ***          *********  * **** * ** ** **

RT-PCR     GCTATATAGGGGGGTGGGTGTTGAGGGAGTCTGGGCAGTCCGTGGGGAACCCCCTTT 237
NOR2&4     GCTATATAGGGGGGTGGGTGTTGAGGGAGTCTGGGCAGTCCGTGGGGAACCCCCTTT 237
F21A14     GCTATATAGGGGGGTGGGTGTTGAGGGAGTCTGGGCAGTCCGTGGG-AACCCCCTTT 236
At3g43160  GCTATATAGGGGG-TGGGTGTTGAGGGAGTCTGGGCAGTCCATGGG-AACCCCCTTA 235
F14M2      GCTATATAGGGGGGTGGGTGTTGAGGGAGTCTGAGCAGTCCGAGGG-AACCCCCTCC 216
           *************  ****************** ******  *** * ********
```

we are unable to distinguish NOR2 and NOR4 rRNA gene transcripts from potential transcripts generated from the truncated single rRNA gene in BAC F21A14, it is likely that the NORs are the source of the transcripts giving rise to the identified smRNAs.

III.4. Discussion

The identification and mapping of smRNA molecules homologous to the rRNA gene sequence revealed a discrete interval overlapping the gene promoter within which 23-24nt siRNAs are abundant. The biogenesis of these siRNAs is dependent on the RNA dependent RNA polymerase RDR2 which is thought to generate the dsRNA precursors sliced by DCL3 into 24nt siRNA duplexes. These observations indicate that the 45S rDNA promoter-homologous smRNAs are generated by the same pathway that generates siRNAs from 5S rRNA *loci* (siR1003) or AtSNI retroelements (Llave *et al,.* 2002; Xi *et al,.* 2004). These so-called heterochromatic siRNAs are involved in mediating DNA methylation at homologous *loci* and, in the case of AtSNI, at least, induce transcriptional gene silencing because higher levels of transcript accumulation are observed in *rdr2* mutant lines (Herr *et al.*, 2005).

Queries of both the ASRP and MPSS *Arabidopsis* databases using the complete sequence of a 45S ribosomal gene unit retrieved a high number of smRNA molecules homologous to the rRNA genes. Taken together, both databases reveal similar distributions of hits across the rRNA gene sequence and the datasets complement each other, with size class information provided by the ASRP data and quantitative information provided by the MPSS data.

The 45S rRNA coding sequence displays a high density of overlapping smRNA hits, frequently displaying sense smRNAs for all the analyzed size classes (21-25nts). This high diversity and density in smRNA molecules is not paralleled in the antisense orientation where smRNA hits are scattered and accumulate to lower levels than their sense orientation counterparts. 45S rRNA coding region smRNAs are quantitatively more abundant than ETS and IGS siRNAs. Curiously, regions near the 3' and 5' ends of the 18S, 5.8S and 25S structural rRNAs seem to be regions for which smRNAs are especially abundant. Coding region smRNAs are not dependent on known RNAi pathways, as evidenced by the fact that there accumulation is unaffected by mutations in any of the four dicers. Northern blot analysis with probes homologous to the 18S (not shown) and 25S coding regions failed to validate the occurrence of homologous smRNA molecules. The long 45S rRNA is transcribed and processed in the nucleolus by excision of transcribed spacers to produce the functional 18S, 5.8S and 25S rRNAs. Transcribed spacers are rapidly degraded and don't accumulate whereas the functional ribosomal RNAs constitute the major fraction of a cell's RNA pool and are the major bands observed upon resolving total RNA in an agarose gel and staining with ethidium bromide. The observed bias

between sense and antisense smRNAs both in size class population and accumulation is indication that these coding region smRNA molecules are likely to be degradation products. This conclusion is supported by the non-involvement of RDR6 in their biogenesis (not shown) and the enrichment of 45S rRNA homologous smRNAs in *rdr2* mutants in which depletion of the 24nt species might allow for relative enrichment of other smRNA species, including miRNAs and 21-22nt siRNAs (Lu *et al.*, 2006). Furthermore, the accumulation of 25S coding sequence homologous 21nt smRNAs in *xrn4* also points to a cytoplasmic rRNA degradation pathway. *XRN4* encodes a homologue of the *S. cerevisae XRNP1p* exoribonuclease involved in RNA degradation (Hsu and Stevens, 1993; Kastenmayer and Green, 2000; Muhlrad and Parker, 1994). In *Arabidopsis sp.* this exoribonuclease localizes in the cytoplasm (Kastenmayer and Green, 2000) where it is involved in mRNA degradation and is implicated in the degradation of 3' end products of miRNA mediated cleavage (Souret *et al.*, 2004). The accumulation of 21nt smRNAs in xrn4 mutants might indicate that loss of function of *XRN4* could trigger an alternative pathway for rRNA turnover or that the exoribonuclease normally act on 21nt smRNA substrates originating from an upstream step of rRNA degradation. Importantly, core promoter and ETS smRNAs accumulation was not significantly affected by this cytoplasmic exoribonuclease. Accumulation of smRNAs in both sense and antisense orientation relative to the 45S rRNA coding sequences points to the existence of dsRNAs, which might be generated by the action of rdRPs, and/or cryptic promoters on both strands. Whole genome approaches to evaluate transcriptional activity in *Arabidopsis* (Lu *et al.*, 2006) have identified 45S rRNA antisense transcripts that could be the source of the identified antisense smRNA molecules.

The IGS region of rRNA genes displays discrete domains of smRNA accumulation in both sense and antisense orientation. One domain corresponds to repeated regions in the IGS upstream of spacer promoters and another is approximately 500bp upstream of the core promoter (note the distribution of smRNA species in figure 3.3a). A distinct domain is localized in the core promoter region where smRNA species overlap the transcription start site. As validated by northern blot hybridization, the 24nt size class rRNA promoter siRNAs are predominant and essentially eliminated in *rdr2* and *dcl3* mutant lines. RDR6 does not participate in their biogenesis nor do DCL1, DCL2 or DCL4 DICER-like enzymes. This observation eliminates the possibility that promoter siRNAs are dependent of the convergence of multiple pathways as is the case for ta-siRNAs (Gasciolli *et al.*, 2005). DCL2 is not required for the generation of the 24nt class although this dicer is apparently involved in the biogenesis of rRNA 22nt siRNAs. Validation of 24nt species identified a discrete interval overlapping the rRNA core promoter and covering the region between -60 to +30 (+1 transcription start site). Read through transcripts were identified overlapping the transcription start site and are likely to originate from rRNA gene units in the NORs and not from other ribosomal sequences elsewhere in the genome.

Spacer promoter sequences are highly diverged from the core promoter except in the -90 to +13 region where near perfect sequence homology exists (not shown) indicating that siRNA species downstream of the transcription start site are core promoter specific and that the former are probably not the source of RNA precursors.

RDR2 and DCL3 were found to colocalize with siRNAs in the nucleolus (Pontes *et al.*, 2006). The siRNA nucleolar processing center is a discrete dot-like domain where siRNAs originating from 45S and 5S rRNA genes as well as AtSNI and *AtCOPIA* retroelements were found to accumulate. Disruption of RDR2 dependent dsRNA precursor synthesis results in delocalization of DCL3 from the nucleolus whereas the opposite is not observed. This is strong evidence that the nucleolar processing center is in fact the nuclear structure where 24nt siRNA biogenesis occurs although it still remains to be verified if RNA precursors are imported as single or dsRNA molecules (Pontes *et al.*, 2006). Core promoter siRNAs validated in this study were also found to localize to the nucleolar processing center and to be essentially eliminated from this nucleolar domain in *rdr2-1* and *dcl3-1* mutant lines, further indicating that they belong to the heterochromatic siRNA class. Probe Atr(+40) failed to identify any significant accumulation of rRNA siRNAs and defined the downstream border of the core promoter siRNA interval. Supporting evidence for the northern blot data was obtained by RNA *in situ* hybridization with a version of this probe designed to identify sense smRNAs. Atr(+40)* hybridized with 45S pre-rRNA transcripts, essentially occupying the entirety of the nucleolus, but the hybridization signal was excluded from the siRNA processing center (Pontes *et al.*, 2006). These observations raise the question of what specifies the RNAs to be used as substrate for siRNA biogenesis. 45S rRNA precursor processing occurs in the nucleolus but does not result in siRNA accumulation. In contrast, siRNAs homologous to the core promoter region are readily identified even though expression levels of transcripts overlapping the transcription start site are much lower. In mammals, IGS derived transcripts are 100 fold less abundant than 45S rRNAs (Mayer *et al.*, 2006). In *Arabidopsis sp.*, S1 nuclease protection assays with DNA probes overlapping the transcription start site hybridizes to 45S rRNA transcripts, an approach that actually determined the exact position of transcription initiation by pol I, and to a lesser extent read-through transcripts at ~10% of the level of transcripts initiating at +1 (Doelling and Pikaard, 1993; 1995). It is likely that NOR promoter siRNAs originate from the processing of a dsRNA whose sense strand originates upstream in the IGS and which hybridizes with an antisense RNA originating from a cryptic promoter downstream of the 45S rRNA transcription start site. It is not clear why the siRNA accumulation is so restricted as RNAs overlapping the core promoter extend beyond the determined interval, but it is possible that it results from cleavage of precursor RNAs before siRNA biogenesis occurs. Active rRNA genes are localized in the nucleolus where there is some evidence that mRNA and tRNA processing also occurs. It can be

speculated that the action of RNA processing machinery, also involved in 45S rRNA maturation, could target and cleave a siRNA dsRNA precursor.

In mammals, epigenetic control of rRNA gene activity is proposed to involve a 150-300nt RNA homologous to the rDNA promoter and generated from a 2kb upstream spacer promoter (Mayer *et al.*, 2006). Association with this non-coding RNA apparently guides NoRC (Nucleolar Remodeling Complex) to the rDNA promoter and triggers H3K9 and H4K20 methylation and HP1 recruitment, hence promoting gene silencing (Santoro *et al.*, 2002; Santoro and Grummt, 2005). Epigenetic regulation of *A. thaliana* rRNA genes is also dependent on DNA methylation and histone post-translational modifications (Chen and Pikaard 1997a; Lawrence *et al.*, 2004). Active genes are essentially demethylated in the promoter region, associated with H3^{trimethyl}K4 and acetylated histones and, importantly RNA polymerase I. Approximately 80% of the rRNA genes are densely methylated and associated with H3^{dimethyl}K9 and are likely to account for the inactive set (Lawrence *et al.*, 2004). In plants, the 24nt siRNA class is involved in sequence specific DNA methylation of homologous regions in the process known as RNA directed DNA methylation (RdDM) (Mette *et al.*, 2000). The siRNA induced epigenetic silencing also involves methylation of H3K9 leading to establishment of a heterochromatic state at the targeted chromatin (Chan *et al.*, 2006; Lippman *et al.*, 2003; Volpe *et al.*, 2003). Targeting of the 45S rRNA promoter and establishment of heterochromatic marks could inhibit pol I binding and transcription initiation. Therefore, it is feasible to raise the hypothesis that the 24nt siRNAs homologous to the promoter could play an important role in rRNA gene epigenetic control.

IV – RNAi mediated chromatin modifications in Arabidopsis thaliana ribosomal genes.

IV.1. The plant specific RNA polymerase IV.

Analysis of the *Arabidopsis* genome sequencing data revealed four genes displaying similarity to the largest and second largest subunits (two genes each) of the nuclear DNA dependent RNA polymerases I, II and III, suggesting the existence of a still uncharacterized fourth class of RNA polymerases in plants (pol IV). Initial evidence pointed to pol IV involvement in the heterochromatic siRNA pathway (Onodera *et al.*, 2005). In order to test this hypothesis smRNA levels were analyzed in distinct T-DNA insertion lines for each of the largest and second largest subunits of this new RNA polymerase. For *NRPD1a*, encoding the largest subunit of Pol IVa, alleles *nrpd1a-3* (SALK_128428) and *nrpd1a-8* (SALK_083051 – not shown) were analyzed. In the case of the second largest subunit, *nrpd2a-1* (SALK_095689 – not shown) and a double mutant *nrpd2a-2/nrpd2b-1* (SALK_046208 and SALK_008535, respectively) were used. Existing evidence points to *NRPD2b* being a non-expressed pseudogene (Onodera *et al.*, 2005) but in order to rule out any possible functional redundancy with *NRPD2a* the double mutant was used for subsequent studies in this work.

Initial evidence for an involvement of pol IV in heterochromatin assembly resulted from the observation that this polymerase was required for DNA methylation of *AtSNI* and 5S rRNA *loci*. As shown in fig. 4.1a, loss of 5S rRNA gene DNA methylation in the CpNpNp sequence context can be evaluated by digestion of genomic DNA with the methylation sensitive endonuclease HaeIII. Except for *dcl3-1*, loss of function mutations in all of the known players of the 24nt siRNA heterochromatic pathway result in demethylation of 5S gene sequences when compared to WT. This loss of methylation phenotype is also observed in mutants defining pol IVa and pol IVb. Loss of function mutations alleles of *DDM1* and *MET1*, which are required for maintenance methylation in CpG and CpNpGp sites but are not involved in 24nt 5S rRNA siRNA biogenesis, display higher levels of DNA methylation. Likewise, *DCL1*, required for miRNA biogenesis, is not involved in 24nt siRNA biogenesis and does not affect DNA methylation at 5S rRNA *loci*. Together, these results confirm the involvement of pol IV in the RdDM pathway (Herr *et al.*, 2005; Kanno *et al.*, 2005; Onodera *et al.*, 2005).

Knock-out of *NRPD1a* and *NRPD2a* in all analyzed T-DNA lines eliminates 24nt siRNAs derived from *AtSNI* retro-elements, 5S rDNA intergenic spacers and 45S rDNA promoter regions (fig. 4.1b). The 21-22nt *AtSNI* siRNAs detected in *nrpd2a-2/nrpd2b-1* homozygous lines was found not to be reproducible (see also fig. 4.2). Evaluation of miRNA accumulation levels (figIV.1b) showed that pol IV is not involved in their biogenesis and thus plays no role in the miRNA pathway. However, pol IVa

Figure 4.1 – A) Analysis of 5S rRNA gene repeat methylation levels in pol IVa (*nrpd1a-3*; *nrpd2a-2/nrpd2b-1*), RdDM pathway, microRNA pathway (*dcl1-7*) and methylation mutants (*ddm1-2; met1-1*). Genomic DNA was digested with Hae III and hybridized to a 5S rRNA gene probe. Southern blot results reflect alterations of DNA methylation levels in a CpNpN sequence context. **B)** Biogenesis of 24nt siRNAs derived from 45S and 5S (probes Atr(-10) and siR1003 respectively) rRNAs and *AtSN1* retroelements is compromised in *nrpd1a-3* and *nrpd2a-2/nrpad2b-1* homozygous mutants (-/-) relative to WT siblings (+/+). Loss of function mutations in pol IVa subunits do not affect miRNA accumulation.

A)

HaeIII - 5S rDNA
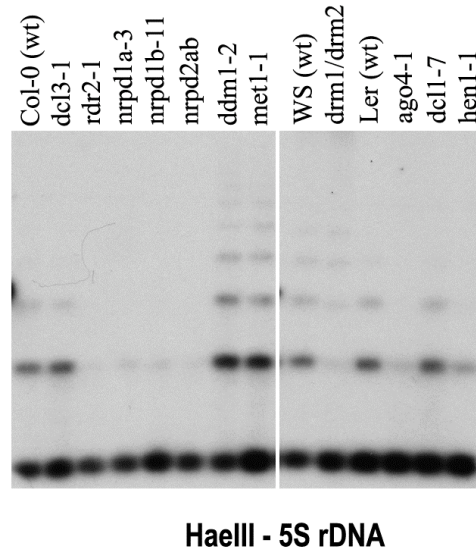
B)
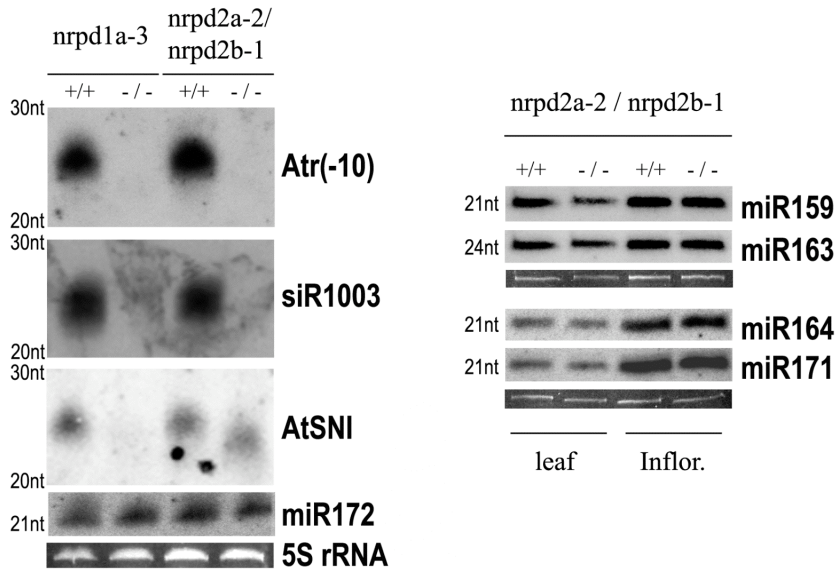
function is not restricted solely to the 24nt siRNA biogenesis pathway as it as been implicated in nat-siRNA production (Borsani *et al.*, 2005).

IV.2. 45S rRNA 24nt siRNA biogenesis is dependent on two distinct pathways.

Biogenesis of 24nt siRNAs homologous to the 45S rRNA IGS and core promoter are dependent on *RDR2* and *DCL3* (see chapter III) and the plant specific pol IVa. These observations are indicative that 45S rRNA siRNAs are products of the endogenous siRNA heterochromatic pathway in *Arabidopsis*. To further characterize their biogenesis, other known players involved in heterochromatic siRNA biogenesis were analyzed and their effects on 45S rRNA siRNAs were compared to *AtSNI*, 5S rRNA, centromere repeats and *AtCOPIA* derived siRNAs (fig. 4.2).

All analyzed 24nt siRNA species were found to be dependent on the activity of DCL3, RDR2, pol IVa and pol IVb confirming previous observations (Herr *et al.*, 2005; Onodera *et al.*, 2005; Pontes *et al.*, 2006, Xie *et al.*, 2004). AGO4 was found to have a mild effect on 45S rRNA core promoter and 5S rRNA siRNAs, contrasting with the strong dependency on this protein for siRNA biogenesis from *AtSNI*, *AtCOPIA* and centromere repeats. siR759, representative of 45S rRNA IGS derived siRNAs, was found to increase its accumulation levels in *ago4-1*. HEN1, involved in smRNA stability (Li *et al.*, 2005), is required for accumulation of all siRNA species analyzed. Typically, siRNA accumulation levels are reduced in *hen1-1* and a size shift is observed, as illustrated by the identification of 25nt siRNA species (compare hybridization patterns on siR759, Atr(-10) and siR1003 between WT Ler and *hen1-1* in fig. 4.2).

DRM2 is the main DNA methyltransferase involved in heterochromatic siRNA biogenesis (Cao *et al.*, 2004). MET1 was found to be required for centromere derived siRNA production (fig. 4.2) (May *et al.*, 2005; this study). 45S rRNA core promoter siRNAs were found to depend on DRM2 for its biogenesis, being essentially eliminated in *drm1/drm2*. In contrast IGS homologous siRNAs (siR759) are dependent on MET1 activity, apparently sharing the same siRNA pathway required for centromere siRNA production. Curiously, the DNA methyltransferases display opposite effects at some *loci*. Specifically, siR1003 accumulates to higher levels in the *met1-1* mutant while siR759 presents the same phenotype in a *drm1/drm2* background.

Finally, effects of chromatin remodeller activities were also evaluated. DRD1 (Kanno *et al.*, 2004) is part of the RdDM pathway in Arabidopsis and found to be involved in 45S rRNA promoter siRNA biogenesis. In contrast, siR759 accumulation is not dependent on this protein but on DDM1 activity (fig. 4.2). This observation is further evidence that 45S rRNA core promoter and IGS siRNAs biogenesis is dependent on two distinct pathways. Core promoter siRNAs are part of the same

Figure 4.2 – Analysis of RNAi pathway members involved in the biogenesis of 45S rRNA homologous siRNAs. 45S rRNA core promoter siRNAs (Atr(-10)) require activity of the same pathway members responsible for 5S rRNA, *AtSNI* and *AtCOPIA* siRNA biogenesis, including DRM2 and DRD1. 45S spacer promoter siRNA (siR759) biogenesis is dependent on DDM1 and MET1 but not on DRM2 and DRD1 activity, as is the also the case for siRNAs homologous to 180bp centromeric repeats. All 24nt siRNA species analyzed by Northern blot are dependent on activity of pol IVa and pol IVb, RDR2, DCL3 and HEN1. EtBr stain of 5S rRNA is shown as loading control.
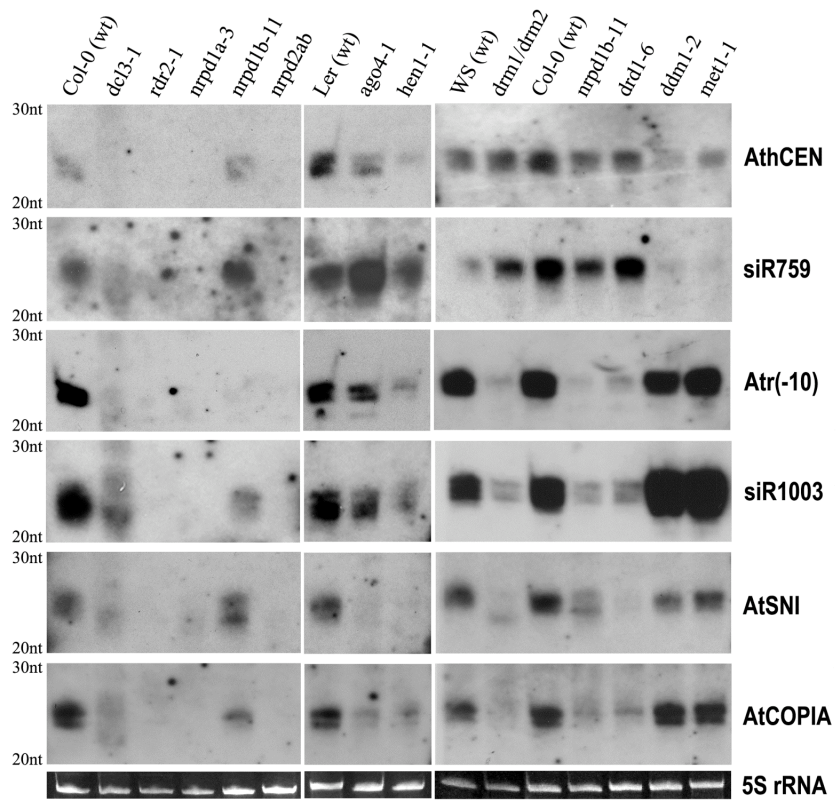
pathway as siR1003, *AtSNI* and *AtCOPIA* siRNAs that are involved in RdDM by a DRM2 and DRD1-dependent pathway. In contrast, IGS derived siRNAs require activity of MET1 and DDM1 for their biogenesis, similar to centromere repeats.
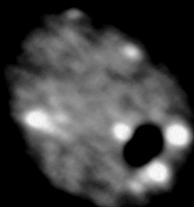
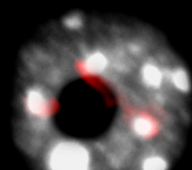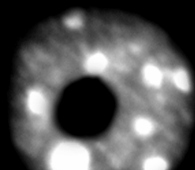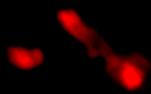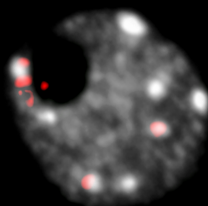## IV.3. The RdDM pathway is involved in interphase organization of NORs.

45S and 5S rRNA interphase organization is altered in an *nrpd2a-2/nrpd2b-1* mutant background (Onodera *et al.*, 2005). DNA *in situ* hybridization was performed in loss of function mutants of the *Arabidopsis* RdDM pathway to further evaluate the role of the endogenous heterochromatic siRNA pathway in the modulation of NOR nuclear organization. Discrete *in situ* hybridization signals were considered and if more than one hybridization signal was identifiable in a cluster each was scored individually. As presented in table 4.1, WT nuclei of the different ecotypes analyzed preferentially display three discrete and compact NOR *foci* (>60% of nuclei examined), perfectly overlapped with the DAPI positive chromocenter, in agreement with previous observations (fig. 4.3a&b) (Fransz *et al.*, 2002). *DDM1* and *MET1* were previously shown not to be involved in NOR interphase organization in *A. thaliana* (Soppe *et al.*, 2002). The obtained results (table 4.1) do not display a significant alteration of the number of *foci* observed in *ddm1-1* and *met1-1* when compared to WT Col-0, although interphase NOR *foci* display a small level of decondensation which is apparent as a fuzzy hybridization signal in the mutant backgrounds (fig. 4.3a). With the exception of *ago4-1*, all other analyzed RdDM pathway mutants displayed significantly altered NOR interphase organization patterns (table IV.1). Amongst the observed phenotypes, FISH signals were frequently diffuse, 45S rRNA *loci* were disrupted into multiple *foci* and, to some degree, association of NORs with chromocenters was lost (fig. 4.3a&b).

Despite significant alterations in the number of *foci*, RdDM mutants do not affect association patterns of NORs in interphase nuclei, as three NOR *foci* were most frequently observed in the mutants, as in wild-type. Therefore, an increase in the number of *foci* can be interpreted as an alteration in the organization of the NORs.

Distribution of frequencies amongst WT ecotypes was found not to be significantly different (χ^2=7.32; d.f.=3; p≤0.2) enabling direct comparison between NOR chromatin organization in RdDM pathway mutants with differ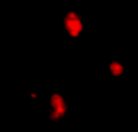ent genetic backgrounds. Interestingly, three different groups of genes were found to mediate significantly different NOR interphase organization patterns. Pol IV subunit mutants and *drm1/drm2* cluster in one group (χ^2<2.79; d.f.=2; p≤1) which is significantly different from a second group composed of *rdr2-1* and *dcl3-1* (*nrpd1a vs dcl3*-1: χ^2=9.74; d.f.=2; p≤0.01 / *rdr2-1 vs dcl3-1*: χ^2=3.14; d.f.=2; p≤1). One explanation for this observation could come from the fact that

Figure 4.3 – A) & B) Analysis of NOR chromatin organization in leaf interphase nuclei of WT and RNAi pathway mutant lines involved in the biogenesis of 45S rRNA homologous siRNAs by fluorescent in situ hybridization (FISH). 45S rDNA probe (red). DNA was counterstained with DAPI (size bar = 5μm).
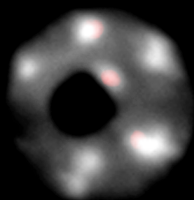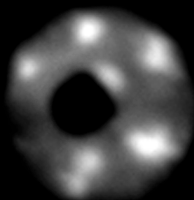
	NORs	DAPI	Merged
WT			
met1			
drm2			
ddm1			
hen1			
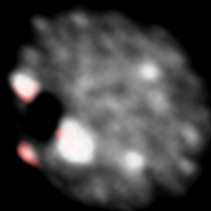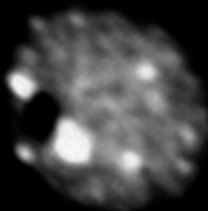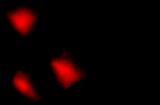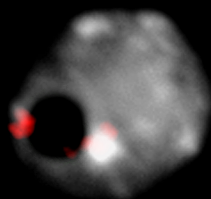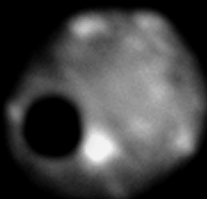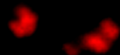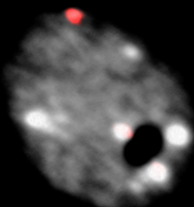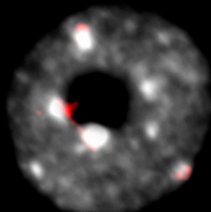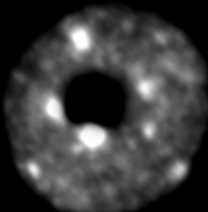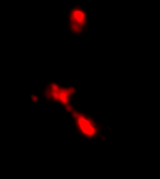
	NORs	DAPI	Merged
nrpd1a			
nrpd2a			
nrpd1b			
rdr2			
dcl3			
ago4			
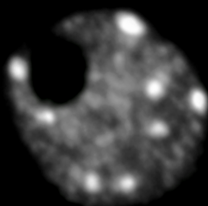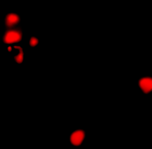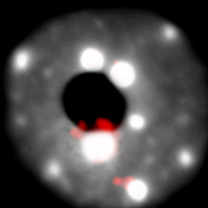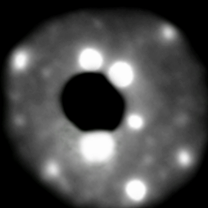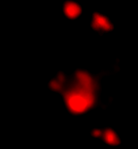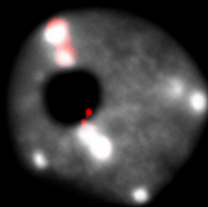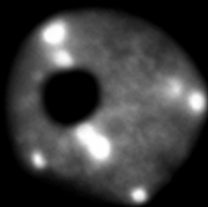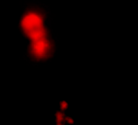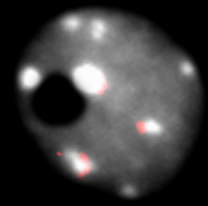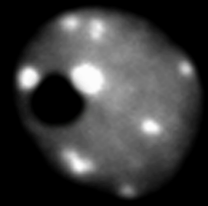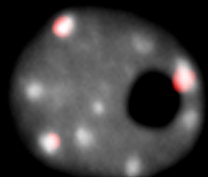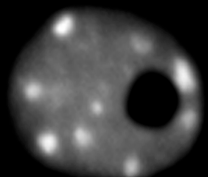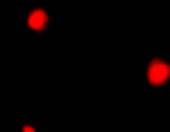
the latter are specific components of the 24nt siRNA heterochromatic pathway whereas the former have been implicated in RNAi based stress response mechanisms. The third group has a single member, *hen1-1*. HEN1, required for smRNA stability, is involved in all known RNAi pathways in *Arabidopsis*. *hen1-1* mutant plants display cumulative defects in development, post transcriptional gene silencing and RdDM which could explain why the distribution amongst the assigned classes is significantly different from the one observed for all other analyzed mutant lines (*nrpd1a vs hen1-1*: χ^2=6.04; d.f.=2; p≤0.05 / *rdr2-1 vs hen1-1*: χ^2=8.99; d.f.=2; p≤0.025) .

Table 4.1 – Frequency (%) of interphase nuclei displaying a distinct number of NOR *foci*, following *in situ* hybridization with an *A. thaliana* 45S rDNA probe. χ^2-test was used to compare NOR patterns observed in mutant genotypes to patterns in WT backgrounds (d.f.=2). The null hypothesis (no difference) was not rejected if p>0.5.

	45S rDNA *foci* number (%)			χ^2	n
	≤ 2	3	≥ 4		
Col-0 (wt)	19	65	16		167
ddm1-2	21	64	15	0.09 (p≤1)	73
met1-1	15	71	14	1.07 (p≤1)	98
nrpd1a-3	17	45	38	19.94 (p≤0.001)	147
nrpd2a	28	41	31	17.22 (p≤0.001)	134
nrpd1b-11	23	38	39	29.44 (p≤0.001)	189
rdr2-1	4	59	37	23.46 (p≤0.001)	97
dcl3-1	8	63	29	11.66 (p≤0.01)	116
WS (wt)	17	75	8		112
drm1/drm2	19	49	32	20.24 (p≤0.001)	103
Ler (wt)	20	61	19		143
ago4-1	15	62	23	1.61 (p≤1)	143
hen1-1	9	41	50	30.59 (p≤0.001)	125

IV.4. 45S rRNA transcript levels are not modulated by the 24nt siRNA pathway.

45S rRNA core promoter siRNAs require for their biogenesis all known members of the RdDM pathway and NOR interphase organization is altered in mutants defining this pathway. In order to evaluate if a correlation exists between these observations and the transcription rate of 45S rRNA genes, multiplex Real Time Quantitative RT-PCR was performed. As presented in figure 4.4a, no

Figure 4.4 – A) Evaluation of transcription levels of 45S rRNA genes in *nrpd1a-3*, *dcl3-1*, *drd1-1*, *nrpd1b-11* and *rdr2-1* mutant lines. Two-step multiplex Real-Time Quantitative RT-PCR was performed with Taqman probes homologous to the 45 rRNA ETS region and Actin (*ACT2*). Relative quantification was normalized to Col-0 (WT). Results presented correspond to three independent experiments and two biological replicates. Error bars depict standard deviation. **B)** Same as in A), dilution series of cDNA input demonstrating that relative quantification data is comparable between independent reactions for the same cDNA input. Results in A) correspond to 1/5 dilution.
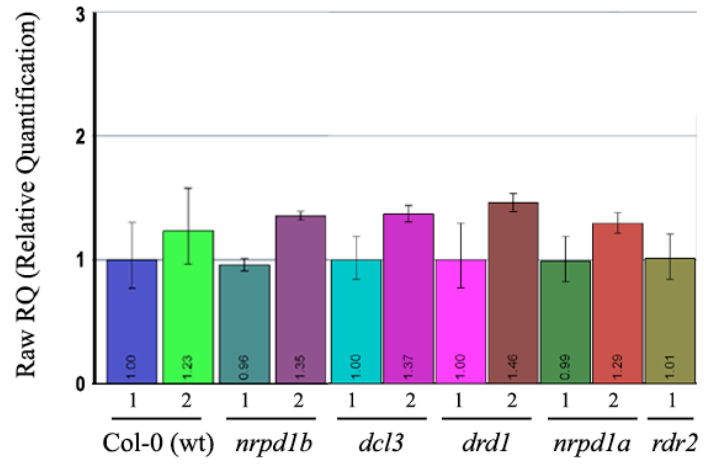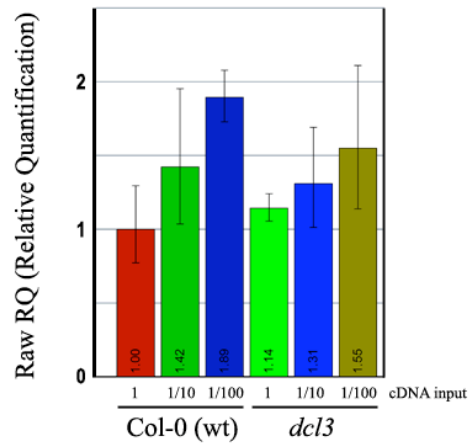
significant changes in 45S rRNA transcript levels were observed between WT and mutant backgrounds.

In order to eliminate the possibility of experimental error due to PCR conditions, a dilution series was performed (fig. 4.4b). Although ratios (RQ) of target sequence (45S rRNA) to endogenous control (*ACT2*) were observed in different experiments, results for WT and *dcl3-1* mutants do not significantly differ for a given template input.

IV.5. Arabidopsis RNPS1 is involved in 24nt siRNA accumulation.

Biogenesis of heterochromatic siRNA was found to occur in the nucleolus (Pontes *et al.*, 2006). Identification and localization of proteins involved in mRNA maturation demonstrated accumulation of an *Arabidopsis* protein similar to *RNPS1* (At1g16610), a member of the exon-joinning complex in mammals, in a discrete nucleolar *focus* (Pendle *et al.*, 2005). *Arabidopsis RNPS1* was found to colocalize with the siRNA nucleolar processing center (fig. 4.5a), indicating a possible link between smRNA precursors and the RNA processing machinery. Additionally, the *Arabidopsis* homolog of the small nuclear RNA binding protein SmD3 (At1g76300) was shown to colocalize in a *Cajal* body with AGO4 (Li *et al.*, 2006), the Argonaute protein involved in the siRNA heterochromatic pathway. SmD3 is a core protein of the small nuclear ribonucleoproteins which are central components of the spliceossome (Khusial *et al.*, 2005). These observations raise the hypothesis that RNA maturation or surveillance mechanisms could be involved in signaling mRNAs to become precursors for siRNA biogenesis.

IV.5.1. Analysis of T-DNA loss of function alleles of *SmD3* and *RNPS1*

In order to evaluate the potential role of RNA processing pathways in smRNA biogenesis, we obtained T-DNA insertion, loss of function *SmD3* and *RNPS1* alleles. SALK_025193 proved to be a *SmD3* null mutant (*smd3-1*) as determined by RT-PCR with primers designed to amplify the 3' end of the mRNA downstream of the T-DNA insertion site (fig. 4.5a). Three different T-DNA lines were analyzed for *RNPS1*. Homozygous mutants were recovered in SALK_004132 (*rnps1-1*) whereas in the other screened transgenic lines only hemizygous or WT siblings were recovered in the progeny of hemizygous plants (SALK_018237: SALK_149423 - not shown). One explanation for these observations is that *rnps1-1* may be a weak allele whereas the other T-DNA insertion lines might result in complete loss of function leading to lethality. RT-PCR identified *rnps1-1* derived transcripts both at the 5' and 3' ends of the gene relative to the T-DNA insertion site, although at lower levels than displayed by the WT allele (fig. 4.5a). Combination of the 5' forward with the 3' reverse primer

readily amplified the expected WT 1.2kb amplicon but no amplification product was observed in the homozygous mutant. The *rnps1-1* allele is therefore transcribed but is likely to be translated into a truncated or altered protein whose function, likely compromised, is still sufficient to allow viability.

IV.5.2. Analysis of 24nt siRNA biogenesis in *rnps1 and smD3 mutants*

Northern blot analysis was performed in order to determine the potential effect of *RNPS1* and *SmD3* genes in the biogenesis of endogenous smRNA species. *SmD3* plays no role in 24nt siRNA accumulation derived from the different *loci* analyzed (fig. 4.5b). Failure to observe accumulation of siRNAs homologous to probe Atr(+40) could be a further indication that RNA splicing is not involved in generating the precursor rRNA gene-derived species to be processed by the siRNA heterochromatic pathway (not shown). One could predict that failure to splice non coding RNAs overlapping the core promoter of 45S rRNA genes at position +17 (relative to the 45S rRNA transcription start site) would extend the siRNA accumulation interval further downstream, beyond its determined 3'border (see Chapter III). These results demonstrate that the putative splicing site located downstream of the core promoter of 45S rRNA genes has no function in siRNA biogenesis at this *locus* (see also Chapter III). In contrast, decreased 24nt siRNA accumulation was observed in *rnps1-1* mutants for 45S and 5S rRNA genes and *AtCOPIA* transposable elements (fig. IV.5b). These genes showed no effect in either miRNA or ta-siRNA biogenesis. Collectively, the results implicate *RNPS1* specifically in 24nt siRNA biogenesis.

IV.5.3. DNA methylation of target *loci* is unaffected in *rnps1-1* mutant line

Heterochromatic siRNAs act as guides for sequence homology-based targeting of effector complexes to DNA sequences. Heterochromatinization of target DNA sites is achieved by DNA methylation and histone modification. Pol IVa is required for siRNA-directed DNA methylation of *AtSN1* and 5S rRNA genes (Herr *et al*, 2005; Onodera *et al.*, 2005). The involvement of RNPS1 in 24nt siRNA accumulation prompted an evaluation of DNA methylation levels at these *loci* (fig. 4.5c). Loss of DNA methylation was readily observed for both DNA sequences in pol IVa mutant lines, confirming previous results. SmD3 and XRN4, shown not to be involved in 24nt siRNA biogenesis (see also Chapter III), had no effect on *AtSN1* or 5S rRNA *loci* methylation levels. Such was also the case for *rnps1-1*, where both methylation in the CpNpNp context in *AtSN1* and CpG and CpNpNp in 5S rRNA genes were essentially unaffected. One possible explanation is that *rnps1-1*, which is probably not a complete loss of function allele, only displays mild effects on siR1003 accumulation levels and that the remaining siRNA pool is still sufficient to guide RdDM.

Figure 4.5 – A) Double immunolocalization and RNA FISH in *A. thaliana* leaf interphase nuclei. A 35S driven RNPS1-FLAG cDNA construct was immunolocalized using anti-FLAG (red). A digoxigenin labeled RNA probe was used for localization of siR1003 (green). DNA was couterstained with DAPI. **B)** Two-step RT-PCR was used to evaluate knock-out of *smd3* and *rnps1* T-DNA lines. For SALK_025193 (*SmD3*), primers were designed to amplify transcription products located downstream of the T-DNA insertion but within the ORF. Genomic DNA was amplified as a control. In the case of SALK_004132 (*RNPS1*), primer pairs to amplify both 5' and 3' ends of the ORF were designed relative to T-DNA insertion. Additionally, forward primer used for 5'-end amplification was combined with reverse primer for 3'-end amplification (T-overlap) in order to identify possible aberrant *RNPS1* RNAs overlapping the T-DNA insertion site. **C)** Northern blot detection of 45S (Atr(-10)), 5S (siR1003) rRNAs and *AtCOPIA* siRNAs, miR159 and siR255 in *smd3-1* and *rnps1-1* mutant backgrounds. EtBr stain of 5S rRNA is shown as loading control. **D)** DNA methylation analysis at *AtSNI* and 5S rRNA *loci* in polIVa mutant lines and *xrn4-5*, *smd3-1* and *rnps1-1*. For *AtSNI*, gDNA was digested with HaeIII (CpNpN) and PCR amplified with primers specific for *AtSNI* or a control gene lacking HaeIII restriction sites. 5S rRNA methylation analysis was performed by Southern blot hybridization of gDNA digested with HaeIII and HpaII (CpG). **E)** *rnps1-1* plants display morphological phenotypes including reduced growth rate and altered leaf morphology when compared to WT Col-0.

A)

α-Flag	siR1003	RNPS1+siR1003	DAPI	Merged

RNPS1-Flag

Col-0

B)

smd3-1: SK_025193

	RT		no RT			Act2		DNA	
	Col-0	smd3	Col-0	smd3	C	Col-0	smd3	Col-0	smd3

1kb
500 bp

rnps1-1: SK_004132

	5' end				3' end				T-overlap				Act2	
	Col-0	rnps1	Col-0	rnps1	Col-0	rnps1	Col-0	rnps1	Col-0	rnps1	Col-0	rnps1	Col-0	rnps1

1.6kb
1kb
500 bp

RT   + + - -   + + - -   + + - -   + +

C)

Col-0 (wt)   rnps1   smd3

30nt
Atr (-10)
20nt

30nt
siR1003
20nt

30nt
AtCOPIA
20nt

21nt   siR255
21nt   miR159

5S rRNA

D)

Col-0 (wt)
0.5 1.0 1.5 nrpd1a-3 nrpd2ab xrn4-5 rnps1 smd3

400bp — AtSNI

200bp — control

Col-0 (wt)
nrpd1a-3
nrpd2ab
xrn4-5
rnps1
smd3
Col-0 (wt)
nrpd1a-3
nrpd2ab
xrn4-5
rnps1
smd3

5S rDNA

HaeIII   HpaII

E)

rnps1 (1 week)   Col-0 (1 week)   0.5cm

rnps1 (3 weeks)   Col-0 (3 weeks)

IV.5.4. *rnps1-1* mutant lines display morphological phenotypes

No visible morphological phenotype was observed for *smd3-1* homozygous plants (not shown). In contrast, *rnps1-1* plants display morphological phenotypes in leaf shape, flowering time and plant development. Seed viability and germination time were unaffected when compared to WT plants. First and second pairs of true leafs display loss of symmetry which results in a curled leaf phenotype (fig. 4.5d). Leaf phenotypes also include serrated margins (most visible at emergence), altered shape and growth rate (fig. 4.5d). *rnps1-1* plants displayed reduced leaf growth rate when compared to WT Col-0 but at 5 weeks after germination both genotypes displayed the same plant diameter (table 4.2). Flowering time is also affected in these mutants (16hours light; 25°C). Three weeks after germination 9/9 Col-0 WT plants had flowered whereas *rnps1-1* inflorescences emerged 5-7 days later. Furthermore, 4/11 mutant plants presented 10 true leafs at flowering whereas all (9/9) Col-0 individual plants displayed 8 true leaves at the same developmental stage, as did the remainder of *rnps1-1* individuals analyzed (7/11).

Table 4.2. – Comparison of *rnps1-1* and Col-0 WT plant diameter (cm) during vegetative growth. The Student T significance test (k=18) of two unknown means and unknown standard deviations was used to compare both populations. 3 weeks after germination *rnps1-1* and Col-0 (WT) plants displayed significantly different plant diameters (τ_{calc}=14.483; p≤0.001). 5 weeks after germination, plant diameter of both genotypes was found not to be significantly different (τ_{calc}=0.649; P=0.736).

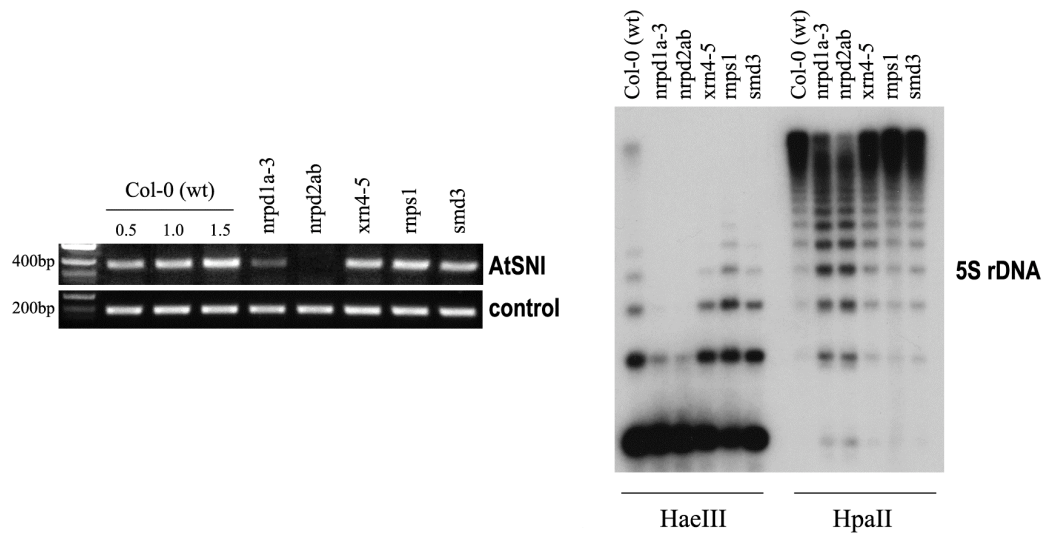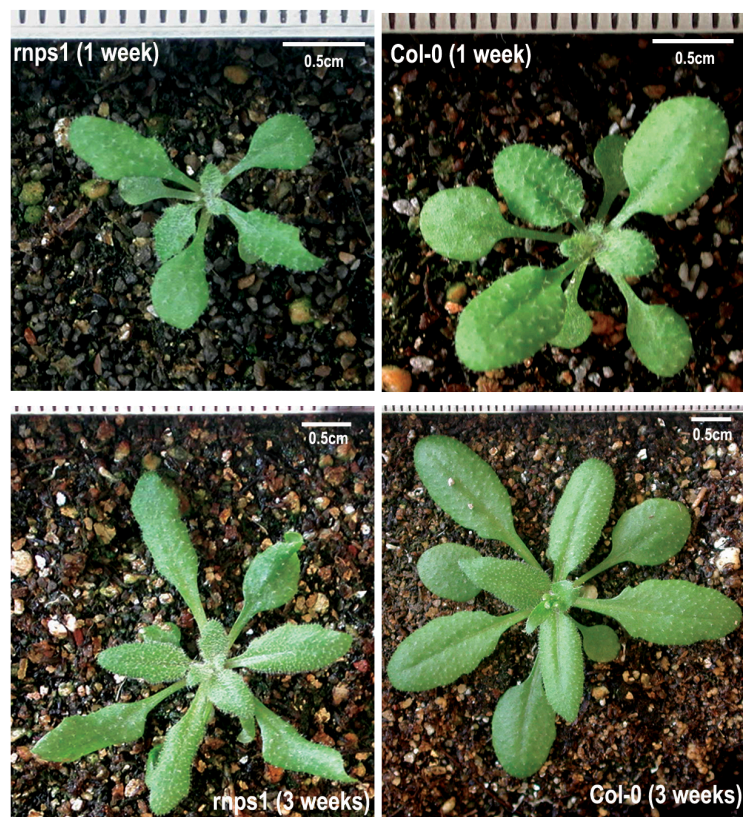| Ø (cm)		*rnps1-1*	Col-0 (WT)
3 weeks	**ū**	3.13	5.02
	σ^2	0.07	0.10
5 weeks	**ū**	8.71	8.61
	σ^2	0.54	0.24
n		11	9

IV.6. Discussion.

Our understanding of the plant specific heterochromatic siRNA pathway has been substantially increased due to recent studies made by a number of research groups in identifying and characterizing key elements of the pathway. As a result of the research conducted in three laboratories (Herr *et al.*, 2005; Onodera *et al.*, 2005, Pontier *et al.*, 2005), and in part to results obtained in this study, a new

player in the form of RNA polymerase IV was identified. *NRPD1a* and *NRPD2a* encode the largest and second largest subunits of pol IVa and loss of function mutations in any of these genes results in total loss of endogenous 24nt siRNA species (fig. 4.1b), a phenotype only paralleled in *rdr2* mutants (Xie *et al.*, 2004). RDR2 was believed to be the primary RNA dependent RNA polymerase responsible for generating RNA precursors to feed the 24nt siRNA pathway (Xie *et al.*, 2004). It has been demonstrated that pol IVa is in fact the upstream factor of the pathway and that RDR2 will act on an RNA template resulting from pol IVa activity (Pontes *et al.*, 2006). However, pol IVa's exact function is still unknown. The largest and second-largest subunits are known to interact with each other (Pontes *et al.*, 2006; Pontier *et al.*, 2005) and association of these subunits is expected to form a functional catalytic domain similar to the conventional DNA dependent RNA polymerases I, II and III. Formation of the catalytic domain using two subunits is a hallmark of DNA-dependent RNA polymerases, which suggests that pol IV should most likely act on a DNA template (Bushnell *et al.*, 2002). In vitro transcription assays have failed to validate this hypothesis (Onodera *et al.*, 2005), which together with the observation that nuclear localization of pol IVa subunits is more affected by RNAse treatment than DNAse treatment (Pontes *et al.*, 2006) points to the possible requirement of an RNA template. In *S. pombe*, pol II transcription is required for targeting of the RITS and RdRC complexes to centromeric regions and establishment of heterochromatic marks (Bühler *et al.*, 2006; Djupedal *et al.*, 2005).

Pol IVa was found to localize essentially to the periphery of the chromocenters and to be required for establishment of heterochromatin (Onodera *et al.*, 2005). The accumulation of the pol IVa complex in discrete nuclear *foci* not only suggests that it acts on very specific *loci* and/or RNA substrates. Pol II, which is responsible for the transcription of all mRNAs, is localized throughout the nucleus but essentially excluded from chromocenters and the nucleolus (Pontes *et al.*, 2006). If pol IVa would act on all RNA species, similar nuclear localization would be expected. Furthermore, RNA substrates for pol IV might result from the transcriptional activity of any of the DNA dependent RNA polymerases. In this regard, it is interesting that pol IVa is involved in the biogenesis of siRNAs derived from 45S rRNA genes transcribed by pol I, 5S rRNA genes and *AtSNI* retroelements transcribed by pol III and *AtCOPIA* transposable elements transcribed by pol II. A feasible possibility is that pol IVa RNA substrates could be non-coding RNAs resulting from the transcription of methylated DNA templates. In support of this hypothesis, DNA methyltransferase activity is required for siRNA biogenesis, with DRM2 involved in the accumulation of some siRNA species and MET1 involved in others (fig. 4.2) (Cao *et al.*, 2003; May *et al.*, 2005; Lippman *et al.*, 2003). Methylated DNA is thought to impair DNA dependent RNA polymerase elongation (Rountree and Selker, 1997; Zilberman *et al.*, 2007). Hence,

stalling or detachment of the transcription complex from its DNA template when finding DNA methylated regions would result in truncated transcripts that might be recognized as substrates for pol IVa activity and might subsequently be processed by the RNAi machinery. The resulting siRNAs, by guiding effector complexes to homologous DNA regions, would increase DNA methylation density leading to the generation of more aberrant RNAs until ultimately transcription would be shut down after consecutive rounds of this circular pathway. In *A. thaliana*, both methylated genes and transposable elements are preferentially located near centromeric regions (Zilberman *et al.*, 2007). One could postulate that preferential localization of pol IVa to chromocenter periphery (Onodera *et al.*, 2005) results from the fact that these regions are likely the major sources of pol IVa RNA templates.

SiRNAs that require MET1 activity for their accumulation apparently do not follow the same rules as DRM-dependent siRNAs. Both siR759, representative of 45S rRNA IGS siRNAs, and 180bp centromeric repeat-derived siRNAs require MET1 and DDM1 for their biogenesis, rather than DRM2. The pathway involving DRM2 and DRD1 is characterized by promoting DNA methylation, as is observed by loss of methylation in target sequences if any of the players is knocked-out in *Arabidopsis*. This is not observed in the case of siR759 or centromeric siRNAs (not shown). *MET1* and *DDM1* were shown to be involved in DNA methylation in CpG and CpNpG sequence contexts. Loss of function mutants for either of these genes results in demethylation of centromeric repeats (Jeddeloh *et al.*, 1999; Kakutani *et al.*, 1995), a phenotype that is not observed in *nrpd1a* and *nrpd2a* mutant lines (Onodera *et al.*, 2005) or as a result of the knock-out of the other players in the siRNA heterochromatic pathway (not shown). In the case of the 45S rRNA IGS siRNA, siR759, the cloned siRNA sequence includes an Aci I restriction site (CGCC). This endonuclease is methylation sensitive and will not cut DNA if any of the cytosines is methylated. Evaluation of the methylation status of this site by Southern blot hybridization indicated that it is fundamentally controlled by MET1 and DDM1 activity. It was only found to be unmethylated in a *ddm1-2* and *met1-1* mutant background whereas all other siRNA heterochromatic pathway mutants displayed WT methylation patterns (not shown). These results, although failing to assign a function to 24nt siRNAs dependent on *MET1* and *DDM1*, support the hypothesis that a DNA methylated template is required for siRNA precursor synthesis. Methylation of DNA in CpG and CpNpG sequence contexts is apparently required for transcription of RNA precursors of 24nt siRNAs. Loss of function mutants in *met1* and *ddm1* display increased levels of transcription of centromeric repeats that, nevertheless, are not recognized by pol IVa as a substrate. This conclusion is supported by the observation that mutations in *rdr2* and *dcl3* were found not to significantly alter transcription of these repeats (May *et al.*, 2005). Furthermore, an increase in 5S

derived siRNAs (siR1003) is observed in *met1-1* and *ddm1-2* mutant lines, where 5S rRNA loci are known to decondense and to be partially excluded from chromocenters (Mathieu *et al.*, 2003). This suggests that loss of methylation leads to increased 5S gene expression, leading to increased siRNA production. Curiously, DNA methylation levels at the CpNpN sequence context, which is presumably due to DRM2, was found to increase in these mutants probably due to increased RdDM (fig. 4.1a). In contrast, CpG and CpNpG methylation is reduced in siRNA pathway mutants, indicating that 5S rRNA *loci* organization requires a balance between the RdDM pathway, involving DRM2, and MET1 and DDM1 activity. This tendency is not observed for all analyzed siRNA species but is observable to a lesser extent in 45S rRNA core promoter derived siRNAs (Atr(-10), fig. 4.2). The opposite is also true, as siR759 accumulates to higher levels in a *drm1/drm2* mutant background, while being dependent on *MET1* for its biogenesis.

The 24nt siRNA heterochromatic pathway in Arabidopsis is thought to be fundamentally involved in transcriptional control of the genome by reestablishing gene transcription patterns after each round of replication. Epigenetic marks of constitutive heterochromatin are transmissible between mother and daughter cells by reestablishment of DNA methylation patterns at CpG and CpNpG symmetrical sequence contexts in a process known to involve DDM1 and MET1 and CMT3 activity. It is also hypothesized that histone post-translational modifications and cell-cycle specific histone isoforms (Ahmad and Henikoff, 2002; Mito *et al.*, 2005), are transmissible and linked to the ability of a daughter cell to reassemble its constitutive heterochromatin fraction. When cells enter S phase DNA replication occurs and nucleosomes are reassembled with both the original as well as newly incorporated histone dimers, partially maintaining heterochromatic marks. 24nt siRNA pathway mutants do not alter the basic nuclear organization as centromere structure is not significantly disturbed (Onodera *et al.*, 2005; Pontes *et al.*, 2006), indicating that this pathway is not required for constitutive heterochromatin formation. Nevertheless, a fraction of 5S and 45S rRNA loci are found to lose association to the chromocenter (Onodera *et al.*, 2005; Pontes *et al.*, 2006; this study). Together with immunolocalization of pol IVa (Onodera *et al.*, 2005) and of pol IVb (Pontes *et al.*, 2006) to the chromocenter periphery it is likely that the 24nt siRNA pathway is required for facultative heterochromatin establishment. Nevertheless, pol IV function is apparently intimately associated with heterochromatic transmissible marks and new evidence strongly supports this idea. Pol IVa was found to be required for the production of >90% of all siRNAs in *A. thaliana* (Zhang *et al.*, 2007). Importantly, the authors found that by crossing *nrpd1a/nrpd1b* and *nrpd2a/nrpd2b* double mutants the resulting F1, carrying a WT copy of each allele, faithfully reestablished 24nt siRNA species found in WT plants. This result clearly indicates that epigenetic marks required for generation of pol IVa RNA

precursors are maintained independently of the functioning of the 24nt siRNA heterochromatic pathway.

The characterization of 45S rRNA gene-derived smRNAs identified 24nt siRNAs clustered in discrete domains of the IGS region. Biogenesis of these smRNA species is dependent on RDR2 and DCL3 indicating that they are part of the endogenous siRNA heterochromatic pathway (see Chapter III). In order to further characterize their biogenesis, accumulation of 24nt siRNAs was evaluated in loss of function mutants of other known players of this pathway (fig. 4.2). Both 45 rRNA spacer (siR759) and core promoter region-homologous siRNAs depend on activity of pol IVa, HENI and pol IVb but otherwise require distinct players for their biogenesis. Spacer promoter siRNAs (siR759) are strongly dependent on DDM1 and MET1 whereas core promoter homologous species (Atr(-10)) require DRM2 and DRD1 for their accumulation. AGO4, which by binding 24nt siRNAs targets effector complexes to homologous DNA sequences (Pontes *et al.*, 2006), also displays different effects on NOR derived siRNAs. Although required for centromere repeats, *AtSN1* and *AtCOPIA* siRNA biogenesis, as well as other *loci* (Qi *et al.*, 2006; Zilberman *et al.*, 2003; 2004), AGO4 shows a weak effect on rRNA derived siRNAs. Both siR1003 (5S rRNA) and 45S gene core promoter siRNA accumulation levels are only slightly reduced in an *ago4-1* mutant background. In the case of siR759 the opposite is observed, with higher accumulation levels in the mutant when compared to WT (fig IV.2). Two different interpretations will necessarily arise from these observations: 1) The *A. thaliana* genome encodes ten Argonaute proteins (Morel *et al.*, 2002) and only four have known function. AGO6 has been recently implicated in 24nt siRNA biogenesis and DNA methylation, being functionally redundant with AGO4 at some *loci* (Zheng *et al.*, 2007). 5S rRNA genes are demethylated in *ago4-1* (fig. IV.1a) (Zilberman *et al.*, 2003) but the same phenotype is not observed in an *ago6-1* background where siR1003 accumulation is also unaffected (Zheng *et al.*, 2007). This observation opens the possibility that other Argonaute proteins are required for an intermediary amplification steps in the case of rRNA *loci*. The resulting 24nt siRNAs loaded in AGO4 and included in the effector complex will finally direct DNA methylation in a two step pathway. 2) Spacer promoter 45S rRNA derived siRNAs were found to be dependent on MET1 and DDM1 activity for their biogenesis. According to the previously discussed hypothesis, establishment of epigenetic marks by these two proteins is required for siRNA precursor synthesis. The increased accumulation of siR759 in *ago4-1* and *drm1/drm2* mutant backgrounds points to the requirement of the RdDM pathway in reducing overall transcription rate of precursor RNAs and reinforcement of the heterochromatic state. Targeting of effector complexes guided by 24nt siRNAs loaded on AGO4 would be necessary for DRM2 recruitment, which by methylating asymmetric cytosine residues would silence transcription units

responsible for RNA precursor synthesis. To substantiate this hypothesis a more thorough evaluation of DNA methylation levels in spacer promoter regions is required.

NOR nuclear organization is dependent on the siRNA heterochromatic pathway (fig. 4.3a&b). Knock-out of *ddm1* and *met1* do not significantly alter the number of 45S rRNA *foci* observed arguing against spacer promoter siRNAs (e.g. siR759) playing a significant role in NOR nuclear organization in *A. thaliana*. The same is true for AGO4 because in an *ago4-1* mutant background NORs displayed WT interphase organization patterns. This Argonaute was also found to not significantly affect core promoter siRNA biogenesis. NOR interphase organization is altered in mutant backgrounds for all other nuclear siRNA pathway members that are required for 45S rRNA core promoter siRNA biogenesis. Nevertheless, pol IVa, pol IVb, DRM2 and HEN1 were shown to be significantly different from RDR2 and DCL3 in their effects. This indicates that other functions performed by the former also impact chromatin organization independently of the RdDM pathway via a still unknown mechanism.

Although 24nt siRNAs homologous to the rRNA genes core promoter are identified and are able to modulate NOR heterochromatin organization in the cell nucleus, no changes in transcription were observed as a result of the knock-out of the RNAi pathway members (fig. 4.4a). 45S rRNA gene transcription rate is controlled not by the number of transcriptionally competent genes in the loci but by the available nuclear pool of pol I complexes engaged in transcription. Studies performed in two strains of *S. cereviseae*, with different copy numbers of rRNA genes, demonstrated that the overall transcription rate of ribosomal RNA was the same despite one strain possessing roughly half the number of gene copies present in the other. The strain with the fewest rRNA genes compensated this deficiency by increasing pol I density on each gene (French *et al.*, 2003).

Establishment of heterochromatic marks has been linked to transcription (Zilberman *et al.*, 2007). Highly transcribed genes are less likely to become methylated suggesting that RNA polymerase transit and density on a given coding sequence impairs silencing complexes from binding chromatin. The Zilberman *et al.* (2007) study was based on wide genome profiling of pol II but some evidence exists that their model is applicable to 45S rRNA gene transcription by pol I. Pre-rRNA synthesis in the cell nucleus is reduced as a result of differentiation and correlates with decreased levels of pol I and an increase in the number of gene units associated with heterochromatic marks (Li *et al.*, 2006). Hence, although the knock-out of members of the endogenous siRNA heterochromatic pathway might result in an increased number of transcriptionally competent 45S rRNA genes, as suggested by interphase nuclear organization, overall pol I transcription rate might be maintained. In the light of current knowledge and of the obtained results one might hypothesize that what is altered is the number of pol I

complexes engaged in transcription of a given gene unit. 45S rRNA genes are highly conserved in sequence which makes it virtually impossible to identify activation of a new set of rRNA genes within a *locus*. The polyploid *A. suecica*, in which nucleolar dominance occurs, is the perfect model system for this analysis and is the subject of chapter V.

Finally, a new potential factor in the 24nt siRNA heterochromatic pathway arises in the form of RNPS1, a constituent of the exon junction complex (EJC) (Le Hir *et al.*, 2000; Lykke-Anderson *et al.*, 2001; Maquat, 2005; Sakashita *et al.*, 2004). The EJC is involved in pre-RNA processing and linked to nonsense mediated decay (NMD) in mammalian cells. NMD is thought to be required not only to degrade transcripts that are abnormal and hence to encode potentially harmful proteins but also to achieve proper levels of gene expression (Maquat, 2005). In *A. thaliana*, RNPS1 was found to localize to the nucleolus in a discrete foci (Pendle *et al.*, 2005), colocalizing with siR1003 (fig. 4.5a), thereby associating this protein with the siRNA nucleolar processing center (Pontes *et al.*, 2006). *rnps1-1*, shown not to be a silent allele (fig. 4.5b), displayed decreased levels of 24nt siRNA accumulation but had no effect in microRNA or ta-siRNA biogenesis (fig. 4.5c). These results, although encouraging and indicating a role for RNPS1 specifically in the 24nt siRNA heterochromatic pathway, did not correlate with DNA demethylation at *bona fide* 24nt siRNA target *loci* (fig. 4.5d). Further work is required to better understand the role of this protein in modulating heterochromatic marks but one could hypothesize that RNPS1 could act on pol IVa generated siRNA RNA precursors and, by its transport to the siRNA nucleolar processing center, participate in its recognition by RDR2 as a substrate.

V - Role of RNAi pathway genes in the modulation of nucleolar dominance in Arabidopsis suecica.
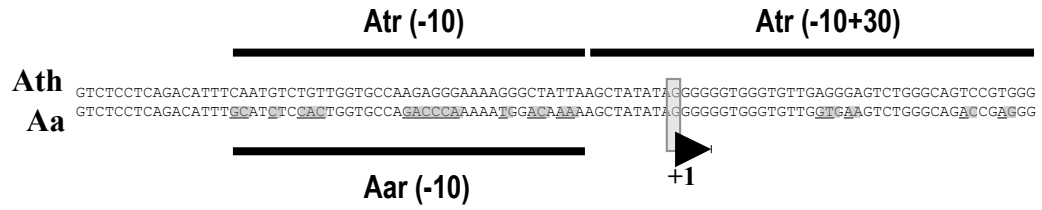
V.1. *Analysis of parental specific 45S rRNA homologous siRNAs- correlations with nucleolar dominance.*

A. suecica LC1 displays nucleolar dominance as a result of the silencing of *A. thaliana* NORs, with 45S rRNA precursors transcribed almost exclusively from *A. arenosa* 45S rRNA genes. Although the 45S rRNA coding sequence of both parental genomes show a high degree of homology they are sufficiently divergent in sequence in the vicinity of the transcription start site to allow the design of parental specific RNA probes (fig. 5.1a). Evaluation of 45S rRNA core promoter siRNAs in *A. arenosa* was performed in the natural accession *A. arenosa* 3651 and demonstrated that in this species the accumulation of promoter homologous siRNAs is restricted to a smaller interval than is the case in *A. thaliana* (fig. 4.1b; see also Chapter III). *A. arenosa* 24nt siRNA species homologous to the region between -40 and -10 (detected with Aar (-10) RNA probe; all sequence coordinates are relative to the 45S rRNA transcription start site) were readily identified. Aar (-10) and the *A. thaliana* specific probe Atr (-10) are species-specific probes, as no cross reactivity was observed in Northern hybridization (fig. 5.1b). The difference between *A. thaliana* and *A. arenosa* siRNA accumulation intervals is that, unlike *A. thaliana*, no significant accumulation of *A. arenosa* 45S rRNA core promoter region siRNAs were observed between positions -10 to +30 (fig. V.1b; Atr(-10+30) probe). Both genomes display highly homologous DNA sequence in the analyzed interval (87.5%), as illustrated in fig. 5.1a, making it unlikely that the Atr (-10+30) probe would not hybridize with *A. arenosa* derived siRNAs. Northern blot hybridization with RNA probes homologous to regions -100 to -60 and +30 to +60 also failed to identify accumulation of 24nt siRNA species in *A. arenosa* 3651 (not shown) indicating the 45S rRNA gene region between -40 to -7 as the main interval of *A. arenosa* 45S rRNA core promoter siRNA accumulation.

Having identified specific 45S rRNA core promoter homologous siRNAs in *A. arenosa*, parental siRNA accumulation levels were evaluated in *A. suecica* in the context of nucleolar dominance. Natural accession *A. suecica* 9502 (Pontes *et al.*, 2003) and transgenic *A. suecica* lines, in an LC1 genomic background, in which *HDA6* (Earley *et al.*, 2006), *HDT1* (Lawrence *et al.*, 2004) or *DRM2* (Preuss and Pikaard, personal communication) are knocked-down through RNAi all display nucleolar dominance disruption. 45S rRNA core promoter siRNA accumulation levels were evaluated in these lines and compared to *A. suecica* LC1, in which nucleolar dominance occurs (fig. 5.1b). It immediately became apparent that siRNA species originating from *A. arenosa* 45S rRNA genes are more abundant than their *A. thaliana* counterparts. Furthermore, *A. thaliana* 45S rRNA siRNAs are strongly reduced in *A. suecica* relative to the diploid *A. thaliana*, an observation not paralleled in the

Figure 5.1 – A) Schematic representation of the DNA sequence of the 45S rRNA core promoter region indicating smRNA probe locations. Polymorphisms in the DNA sequence between *A. thaliana* and *A. arenosa* derived rRNA genes are highlighted and enabled the design of species-specific RNA probes Atr(-10) and Aar(-10), respectively. **B)** Northern blot hybridization of low molecular weight (<500nt) leaf RNA. Species-specificity of RNA probes was verified as no significant cross hybridization was observed between parental lines *A. arenosa* 3651 (Aa 3651) and *A. thaliana* No-0 (Ath No-0). Accumulation of 45S rRNA promoter homologous siRNAs was evaluated in *A. suecica* LC1 (As LC1 (WT)), in which nucleolar dominance occurs, and compared to *A. suecica* lines in which nucleolar dominance is disrupted. Knock-out of *hda6* (*axe1-5*) leads to increased levels of 45S rRNA homologous smRNA accumulation in *A. thaliana*, a phenotype replicated in *A. suecica* by RNAi mediated knock down of HDA6 mRNA. miR163 and EtBr stained 5S rRNA are shown as loading controls. **C)** S1 nuclease protection assay demonstrating that nucleolar dominance is disrupted in the analyzed *A. suecica* lines but observed in wild-type (non-transgenic) *A. suecica* LC1.

A)

Atr (-10)   Atr (-10+30)

Ath GTCTCCTCAGACATTTCAATGTCTGTTGGTGCCAAGAGGGAAAAGGGCTATTAAGCTATATAGGGGGTGGGTGTTGAGGGAGTCTGGGCAGTCCGTGGG

Aa GTCTCCTCAGACATTTGCATGTCCACTGGTGCCAGACCCAAAAATGGACGAAAAGCTATATAGGGGGTGGGTGTTGGTGAAGTCTGGGCAGACCGAGGG
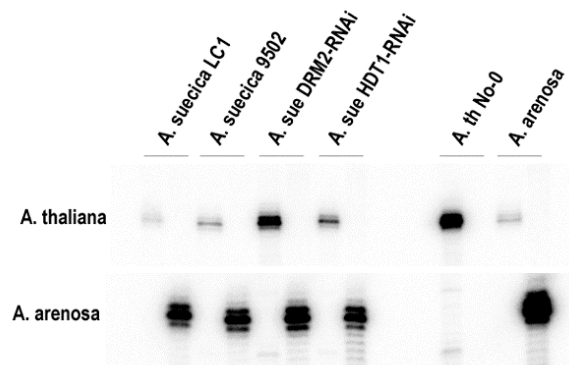
Aar (-10)

+1

B)
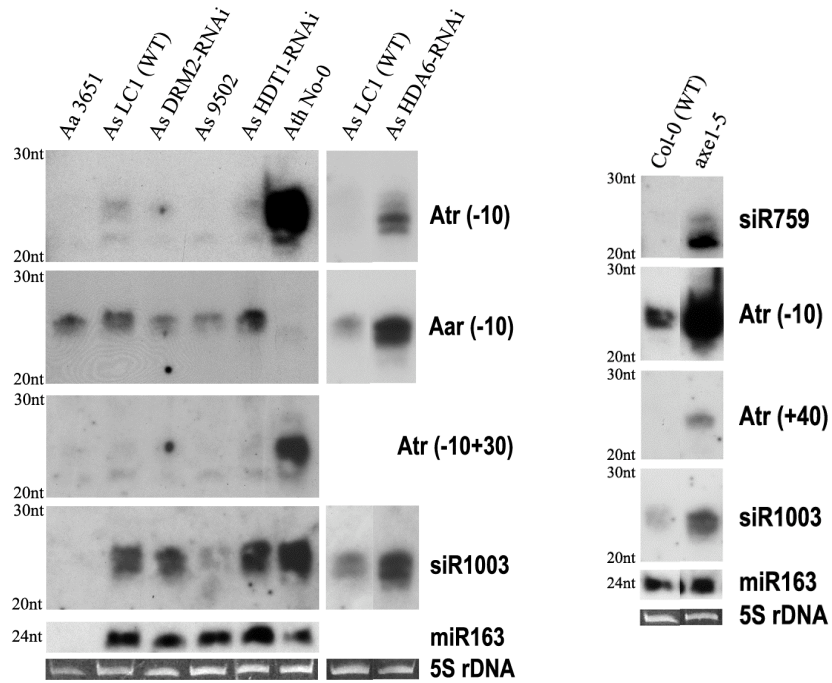
C)

case of 5S rRNA homologous siRNAs (siR1003). Curiously, siR1003 does not accumulate in *A. arenosa* 3651 despite being homologous to the 5S rRNA gene repeats (not shown) and miR163 was also not identified in the tetraploid.

A. suecica 9502 displays decreased accumulation of all analyzed siRNAs relative to LC1 (fig. 5.1b), a phenotype that is unique amongst all other *A. suecica* lines studied. DRM2-RNAi lines showed decreased accumulation levels of 45S rRNA derived siRNAs but such was not the case for siR1003 (5S rRNA). HDT1-RNAi lines displayed wild-type siRNA accumulation levels indicating that this putative histone deacetylase (Pandey *et al.*, 2002) is not involved in 24nt siRNA biogenesis. Finally, HDA6-RNAi lines display a substantial increase in accumulation levels of rRNA-derived siRNAs independent of their parental origin. In order to further evaluate the effect of HDA6 loss of function on rRNA homologous siRNAs, the null allele *axe1-5*, in an *A. thaliana* genomic background, was analyzed. As presented in fig. 5.1b, the same increases in rRNA siRNA accumulation levels relative to WT were readily observed. Preferential accumulation of a new 21nt siRNA species homologous to the siR759 RNA probe was also identified (see Chapter IV). Also significant is the identification by Northern hybridization of a 24nt siRNA with probe Atr (+40) indicating that in a *axe1-5* mutant background the discrete interval of 24nt siRNA accumulation at the 45S rRNA core promoter region is altered (see also Chapter III).

To validate the performed siRNA analysis presented in fig. 5.1b in the context of nucleolar dominance, high molecular weight RNA fractions (>500nt), extracted from the same leaf lysate used for smRNA blots (<500nt), were used in S1 protection assays (fig. 5.1c). Identification of *A. thaliana* derived 45S rRNA precursor transcripts in *A. suecica* 9502, HDT1-RNAi and DRM2-RNAi indicates that nucleolar dominance is in fact disrupted in these lines.

V.2. The 24nt heterochromatic siRNA pathway and nucleolar dominance in A. suecica.

As described in Chapter II, WT *A. suecica* LC1 was transformed in order to achieve knock-down of genes known to be involved in the 24nt siRNA heterochromatic pathway in *A. thaliana*. Progeny (T_1) of transformed WT plants (T_0) were grown on selective medium and viable T_1 transformants were screened for knock-down levels of the target genes as well as effects on DNA methylation levels at *AtSN1 loci*, a *bona fide* target of RdDM (Herr *et al.*, 2005; Onodera *et al.*, 2005; Xie *et al.*, 2004). As expected, RNAi mediated knock-down efficiency varied amongst T_1 individuals (fig 5.2-6a) but in general was found to correlate with the observed levels in *AtSN1* DNA methylation (fig 5.2-6b). RNAi transformants were then analyzed for nucleolar dominance disruption (fig 5.2-6c,d) and rRNA

homologous 24nt siRNA accumulation levels (fig 5.2-6e). Evaluation of the transcription status of parental 45S rRNA genes was performed by using an S1 nuclease protection assay (Chen *et al.*, 1997a; Doelling *et al.*, 1993) and CAPS (Lewis and Pikaard, 2001).

V.2.1. *A. suecica* LC1 NRPD2-RNAi

Analysis of NRPD2a mRNA levels in NRPD2-RNAi T_1 transformants revealed that only individuals *11* and *44* did not show significant knock-down of the target mRNA, and for individual *44* no alterations in *AtSN1* DNA methylation levels were observed when compared to WT (fig. 5.2a;b). Individuals *21* and *31* display an approximately 0.5 fold reduction in NRPD2a transcript accumulation, which translates into a slight decrease in *AtSN1 loci* methylation. In contrast, individual *22*, although displaying target gene knock-down to a similar level as in individual *21*, displays a strong reduction in DNA methylation levels. Transformants *41-43* show a substantial knock-down of NRPD2a which correlates well with the degree of and demethylation of *AtSN1* DNA sequences.

NRPD2-RNAi T1 transformants typically showed a mild degree of reactivation of A. thaliana 45S rRNA genes when compared to WT LC1, with transformant *41* being an exception that showed strong reactivation of *A. thaliana* rRNA genes in both assays (fig. 5.2c;d). Taking into account the NRPD2 mRNA knock-down levels and the effect on *AtSN1 loci* DNA methylation one could expect individuals *41*, *42* and *43* would display similar levels of *A. thaliana* 45S rRNA reactivation, but this was not observed. Intriguing is the observation of disruption of nucleolar dominance in transformant *11* which does not display significant NRPD2 mRNA knock-down. A possible interpretation for this observation is that it is due to an artifact derived of transgene(s) integration site or that RNAi may interfere with translation rather than simply cause mRNA target degradation.

NRPD2-RNAi line *31* displays a moderate knock-down of the target mRNA, *AtSN1* demethylation and does not disrupt nucleolar dominance. As shown in figure 5.2e, 45S and 5S rRNA siRNA biogenesis is not compromised relative to wild type LC1 in this transformant. In contrast *41*, *42* and *43* show a strong decrease in 24nt siRNA accumulation levels correlating with the observed *NRPD2a* knock-down levels. Transformant *41* shows the strongest effect being also the one where *A. thaliana* 45S rRNA transcripts are more abundant. Individuals *42* and *43* display an intermediate effect in siRNA accumulation and a low level of *A. thaliana* 45S rRNA gene reactivation.

V.2.2. *A. suecica* LC1 RDR2-RNAi

A. suecica RDR2-RNAi T_1 plants show diverse RNAi efficiency ranging from no visible effect (*21* and *44*) to severe knock-down of endogenous RDR2 transcript levels (*41-43*). Effects on *AtSN1* DNA methylation ranged from essentially not observable (*44* and *45*) to a significant decrease obtained for

Figure 5.2 – Screen of RNAi knock-down efficiency, 24nt siRNA heterochromatic pathway targets and nucleolar dominance disruption in *A. suecica* NRPD2-RNAi T_1 transformants (numbered). **A)** Semi-quantitative RT-PCR with NRPD2a specific primers was used to verify knock-down levels of target transcripts. **B)** DNA methylation analysis at *AtSNI loci* in *A. suecica*. Semi-quantitative PCR of gDNA digested with HaeIII (CpNpN) was performed with primers specific for *AtSNI loci* or a control gene lacking HaeIII restriction sites. **C, D)** Evaluation of nucleolar dominance disruption as a result of RNAi mediated knock-down of NRPD2a mRNA in *A. suecica* LC1 by C) S1 nuclease protection assay and D) CAPS. **E)** Analysis of 45S (Atr(-10); Aar(-10)) and 5S (siR1003) rRNA homologous smRNAs, miR163 and siR255 (ta-siRNA) accumulation levels in RNAi individual T_1 transformants. EtBr stain of 5S rRNA is shown as RNA loading control.

A)

B)

C)

D)

E)
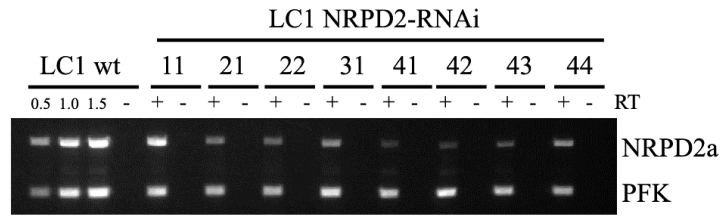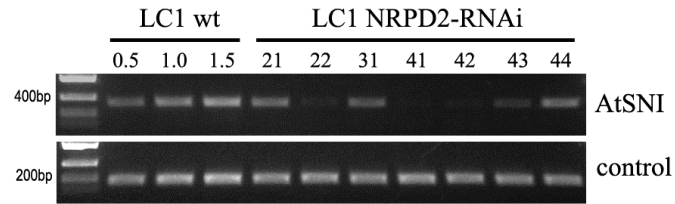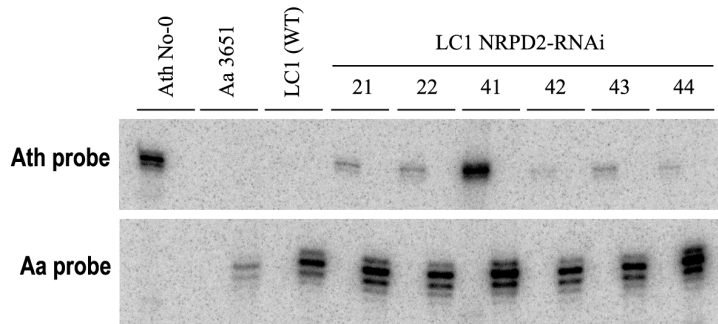
Figure 5.3 – Screen of RNAi knock-down efficiency, 24nt siRNA heterochromatic pathway targets and nucleolar dominance disruption in *A. suecica* RDR2-RNAi T$_1$ transformants (numbered). **A)** Semi-quantitative RT-PCR with RDR2 specific primers was used to verify knock-down level of target transcript. **B)** DNA methylation analysis at *AtSNI loci* in *A. suecica*. Semi-quantitative PCR of gDNA digested with HaeIII (CpNpN) was performed with primers specific for *AtSNI loci* or a control gene lacking HaeIII restriction sites. **C, D)** Evaluation of nucleolar dominance disruption as a result of RNAi mediated knock-down of RDR2 mRNA in *A. suecica* LC1 by C) S1 nuclease protection assay and D) CAPS. **E)** Analysis of 45S rRNA (Atr(-10); Aar(-10)) homologous smRNAs, miR163 and siR255 (ta-siRNA) accumulation levels in RNAi individual T$_1$ transformants. EtBr stained 5S rRNA is shown as RNA loading control.

A)

	LC1 wt			LC1 RDR2-RNAi							
	0.5	1.0	1.5	21	22	31	41	42	43	44	45

RDR2

PFK

noRT (RDR2)

B)

	LC1 wt			LC1 RDR2-RNAi							
	0.5	1.0	1.5	22	23	31	41	42	43	44	45

500bp — AtSNI

500bp — control

C)

Ath No-0
Aa 3651
LC1 (WT)

LC1 RDR2-RNAi
31 41 42 43 44 45

Ath probe

Aa probe

D)

A. thaliana
LC1 wt 0.5 1.0 1.5
DRM2-RNAi
Aa 3651

LC1 RDR2-RNAi
21 22 23 31 41 42 43 44 45

E)

LC1 (WT)

LC1 RDR2-RNAi
42 43 44 45

30nt
20nt Atr (-10)

30nt
20nt Aar (-10)

21nt siR255

24nt miR163

5S rRNA

individual *43*. Transformants *41* and *42*, despite displaying around 0.5 fold reduction in RDR2 mRNA levels, show a limited effect on *AtSNI* methylation when compared to individual *43* (fig 5.3a;b).

Individual plant *43*, whose significant *RDR2* knock-down correlates with *AtSNI loci* demethylation, also displays *A. thaliana* derived 45S rRNA gene reactivation (fig. 5.3c;d). Other analyzed individual transformants show reduced to no reactivation of *A. thaliana* genes despite the fact that some display efficient target gene knock-down which in turn is sufficient to, at a limited degree, impair *AtSNI* DNA methylation. The observed results suggest that only in the case of transformant *43* is the *RDR2* knock-down level sufficient to compromise gene function to a significant level.

RDR2 is required for 24nt siRNA biogenesis in *A. thaliana* and knock-out of this gene results in elimination of all 24nt siRNA species (Xie *et al.*, 2004). None of the analyzed transformants replicated this phenotype in *A. suecica* RDR2-RNAi lines indicating that, despite the knock-down of the target gene, the remaining RDR2 RNA pool is still sufficient to mediate 24nt biogenesis. Nevertheless, transformant *43* displays a significant reduction in rRNA siRNA accumulation, especially in the case of *A. thaliana* derived species (fig. 5.3e).

V.2.3. *A. suecica* LC1 DCL3-RNAi

Although only four transgenic DCL3-RNAi T$_1$ transformants were recovered from the selection screen all display significant decreases in DCL3 mRNA accumulation when compared to WT LC1 (fig. 5.4a). Individual *41* is apparently the one with the least efficient knock-down. *DCL3* does not severely affect DNA methylation at *AtSNI loci* in *A. thaliana* (Onodera *et al.*, 2005) and, with the exception of transformant *32*, no significant alterations in DNA methylation at these *loci* were observed in T$_1$ plants (fig. 5.4b).

Analysis of parental specific 45S rRNA gene expression readily identified A. thaliana derived transcripts in all analyzed transformants (fig. 5.4c;d), implicating DCL3 in the modulation of nucleolar dominance.

A *dcl3-1* like siRNA hybridization pattern (Pontes *et al.*, 2006; Xie *et al.*, 2004; this study) is observed in individual *32* but not in the other analyzed transformants. Nevertheless, a decrease in 24nt siRNA accumulation levels is observed for individual*s 31* and 33 relative to WT LC1 (fig. 5.4e). Transformant *41*, shown to have the least efficient DCL3 knock-down, does not display any effect on siRNA accumulation. Both miR163 and siR255 accumulation levels, which in *A. thaliana* are dependent on DCL1 and DCL1+DCL4 activity, respectively, were not affected in DCL3-RNAi transformants relative to WT LC1.

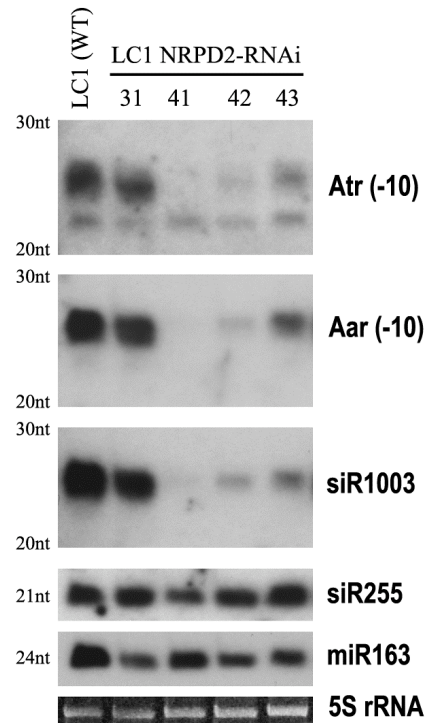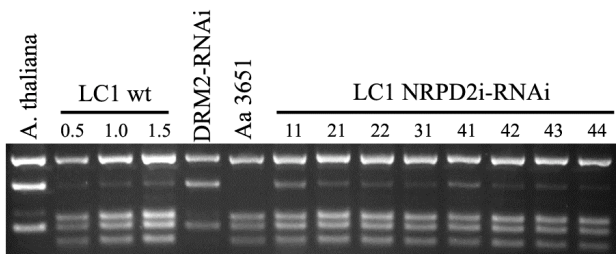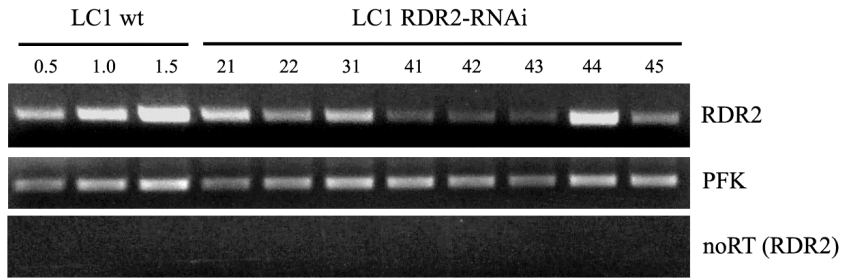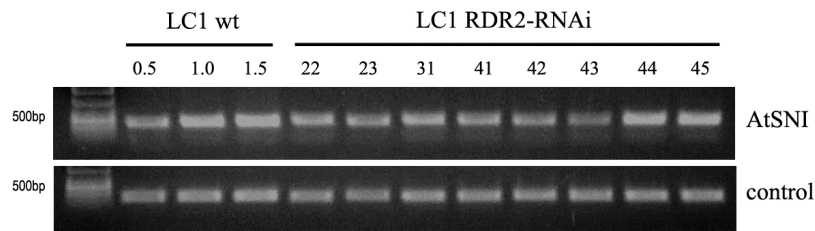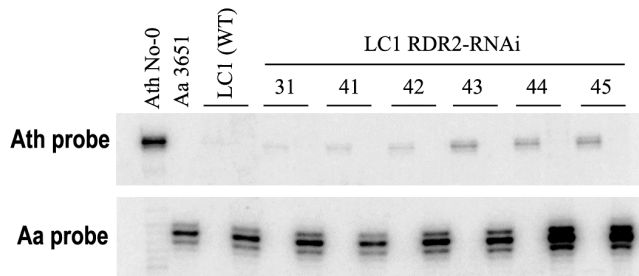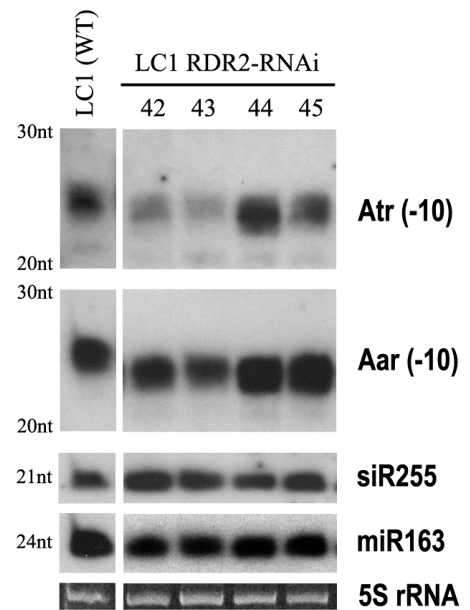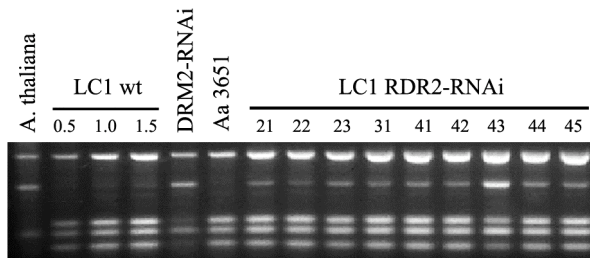**Figure 5.4 –** Screen of RNAi knock-down efficiency, 24nt siRNA heterochromatic pathway and nucleolar dominance disruption in *A. suecica* DCL3-RNAi T_1 transformants (numbered). **A)** Semi-quantitative RT-PCR with DCL3 specific primers was used to verify knock-down level of target transcript. **B)** DNA methylation analysis at *AtSNI loci* in *A. suecica*. Semi-quantitative PCR of gDNA digested with HaeIII (CpNpN) was performed with primers specific for *AtSNI loci* or a control gene lacking HaeIII restriction sites. **C, D)** Evaluation of nucleolar dominance disruption as a result of RNAi mediated knock-down of DCL3 mRNA in *A. suecica* LC1 by C) S1 nuclease protection assay and D) CAPS. **E)** Analysis of 45S (Atr(-10); Aar(-10)) and 5S (siR1003) rRNA homologous smRNAs, miR163 and siR255 (ta-siRNA) accumulation levels in RNAi individual T_1 transformants. EtBr stained 5S rRNA is shown as RNA loading control.

A)

	LC1 wt			LC1 DCL3-RNAi			
	0.5	1.0	1.5	31	32	33	41

DCL3

PFK

noRT (DCL3)

B)

	LC1 wt			LC1 DCL3-RNAi			
	0.5	1.0	1.5	31	32	33	41

400bp — AtSNI

200bp — control

C)

Ath No-0   Aa 3651   LC1 (WT)   LC1 DCL3-RNAi
31   32   33   41

Ath probe

Aa probe

D)

A. thaliana   Aa 3651   LC1 wt   DRM2-RNAi   LC1 DCL3-RNAi
0.5   1.0   1.5   31   32   33   41

E)

LC1 (WT)   LC1 DCL3-RNAi
31   32   33   41

30nt — 20nt   Atr (-10)

30nt — 20nt   Aar (-10)

30nt — 20nt   siR1003

21nt   siR255

24nt   miR163

5S rRNA
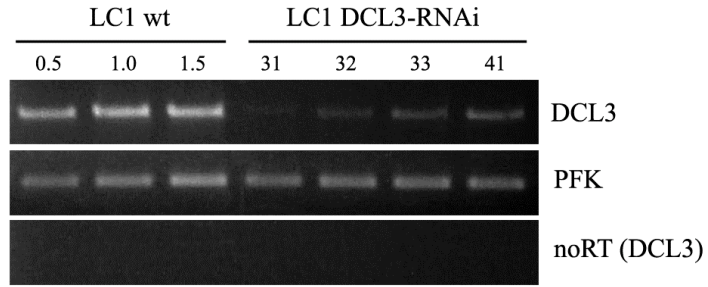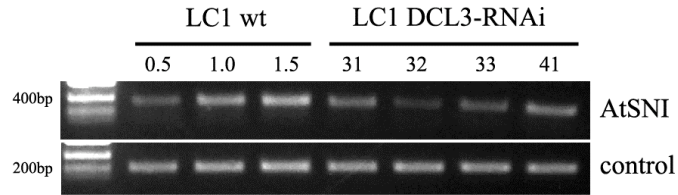
V.2.4. *A. suecica* LC1 HEN1-RNAi

HEN1-RNAi T_1 transformants recovered from the selective screen demonstrate significant degrees of HEN1 mRNA knock-down (fig. 5.5a). *HEN1* has no severe effect in *AtSN1* methylation levels in the diploid *A. thaliana* (Onodera *et al.*, 2005), as is also observed in fig. 5.5b, although some decrease is apparent in individuals *14* and *42* when compared to WT LC1.
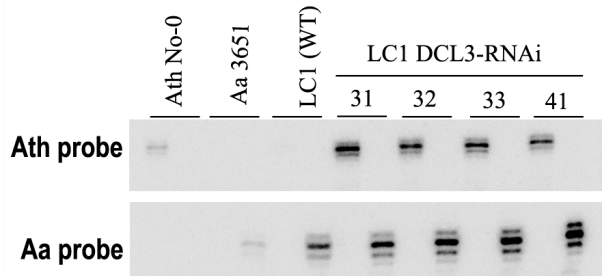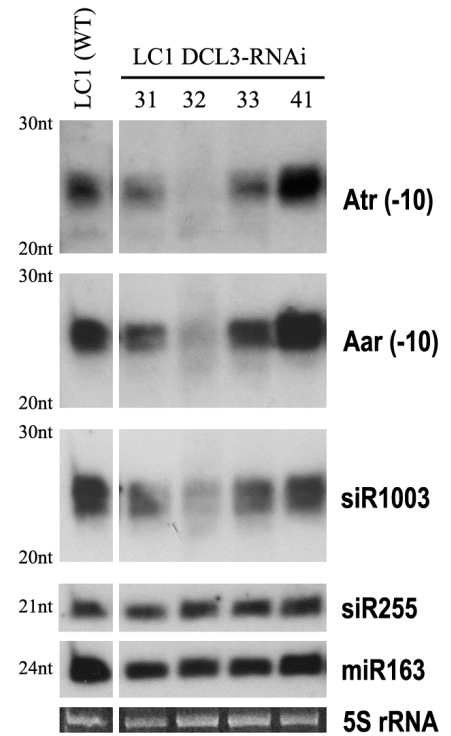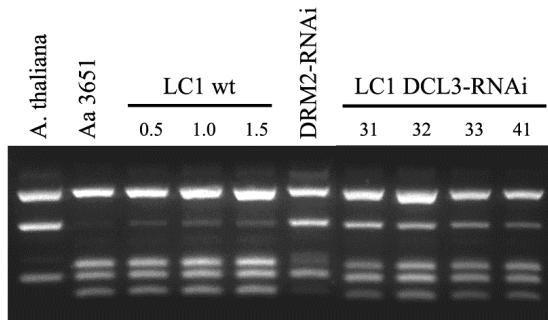
Except in the case of individual *22*, all other analyzed HEN1-RNAi T_1 transformants display reactivation of *A. thaliana* 45S rRNA genes, with the strongest reactivation of the underdominant rRNA class observed in transformants *12*, *14* and *21* (fig. 5.5c;d).

SmRNA analysis in HEN1-RNAi T_1 transformants demonstrated the efficiency of the RNAi mediated knock-down of this gene, as microRNA accumulation was severely impaired. Nevertheless, only transformant *14* displays equivalent effects on both 24nt siRNAs and tasiRNAs when compared to WT (fig. 5.5e). This result is somewhat unexpected as *hen1* mutant lines in an *A. thaliana* genomic background display reduced accumulation levels of all smRNA species independently of their biogenesis pathway (Xie et al., 2004; Yang *et al.*, 2006; in this study).

V.2.5. *A. suecica* LC1 AGO4-RNAi

AGO4 is a major player in RdDM at *AtSN1 loci* in *A. thaliana*, with both homologous siRNA biogenesis and DNA methylation strongly dependent on AGO4 activity (Herr *et al.*, 2005; Yigit *et al.*, 2006; Zilberman *et al.*, 2003). Such also seems to be the case in *A. suecica* as a strong correlation was observed between knock-down levels of AGO4 mRNA and *AtSN1* methylation levels in the analyzed AGO-RNAi transformants (fig. 5.6a;b). Transformants *14* and *21* which display low levels of target gene knockdown were also found to essentially display WT *AtSN1* methylation levels. In contrast, all other T_1 individuals show strong gene knock-down levels correlated with severe reduction of DNA methylation (fig. 5.6a;b).

Despite an apparent highly efficient knock-down of AGO4 mRNA only transformants *21* and *31* display disruption of nucleolar dominance (fig. 5.6c;d). The case of transformant *21* is unexpected because if low levels of AGO4 mRNA knock-down were sufficient to reactivate *A. thaliana* 45S rRNA genes, all transformants shown to efficiently impair target mRNA accumulation should display the same if not stronger reactivation phenotype.

LC1 AGO4-RNAi individual *31* displays disruption of nucleolar dominance and was also found to be the transformant in which 24nt siRNA accumulation levels were most severely affected (fig. 5.6e). Although only a mild effect is observed in the case of rRNA derived siRNAs, agreeing with previous

Figure 5.5 – Screen of RNAi knock-down efficiency, 24nt siRNA heterochromatic pathway and nucleolar dominance disruption in *A. suecica* HEN1-RNAi T_1 transformants (numbered). **A)** Semi-quantitative RT-PCR with HEN1 specific primers was used to verify knock-down level of target transcript. **B)** DNA methylation analysis at *AtSNI loci* in *A. suecica*. Semi-quantitative PCR of gDNA digested with HaeIII (CpNpN) was performed with primers specific for *AtSNI loci* or a control gene lacking HaeIII restriction sites. **C, D)** Evaluation of nucleolar dominance disruption as a result of RNAi mediated knock-down of HEN1 mRNA in *A. suecica* LC1 by C) S1 nuclease protection assay and D) CAPS. **E)** Analysis of 45S (Atr(-10); Aar(-10)) and 5S (siR1003) rRNA homologous smRNAs, miR163 and siR255 (ta-siRNA) accumulation levels in RNAi individual T_1 transformants. EtBr stained 5S rRNA is shown as RNA loading control.

A)

LC1 HEN1-RNAi

	LC1 wt				11	12	14	21	22	42	43	51

0.5 1.0 1.5 - C | + - | + - | + - | + - | + - | + - | + - | + - RT

HEN1

PFK

B)

LC1 wt | LC1 HEN1-RNAi

0.5 1.0 1.5 | 11 12 14 21 22 42 43

400bp — AtSNI

200bp — control

C)

Ath No-0 | Aa3651 | LC1 (WT) | LC1 HEN1-RNAi

12 14 21 22 42 43

Ath probe

Aa probe

D)

A. thaliana | LC1 wt | DRM2-RNAi | Aa 3651 | LC1 HEN1-RNAi

0.5 1.0 1.5 | | | 11 12 14 21 22 42 43 51

E)

LC1 (WT) | LC1 HEN1-RNAi

12 14 21 42

30nt
Atr (-10)
20nt

30nt
Aar (-10)
20nt

30nt
siR1003
20nt

21nt — siR255

24nt — miR163

5S rRNA

Figure 5.6 – Screen of RNAi knock-down efficiency, 24nt siRNA heterochromatic pathway and nucleolar dominance disruption in *A. suecica* AGO4-RNAi T_1 transformants (numbered). **A)** Semi-quantitative RT-PCR with AGO4 specific primers was used to verify knock-down levels of target transcript. **B)** DNA methylation analysis at *AtSNI loci* in *A. suecica*. Semi-quantitative PCR of gDNA digested with HaeIII (CpNpN) was performed with primers specific for *AtSNI loci* or a control gene lacking HaeIII restriction sites. **C, D)** Evaluation of nucleolar dominance disruption as a result of RNAi mediated knock-down of AGO4 mRNA in *A. suecica* LC1 by C) S1 nuclease protection assay and D) CAPS. **E)** Analysis of 45S rRNA (Atr(-10); Aar(-10)), *AtCOPIA* homologous smRNAs, miR163 and siR255 (ta-siRNA) accumulation levels in RNAi individual T_1 transformants. EtBr stained 5S rRNA is shown as RNA loading control.
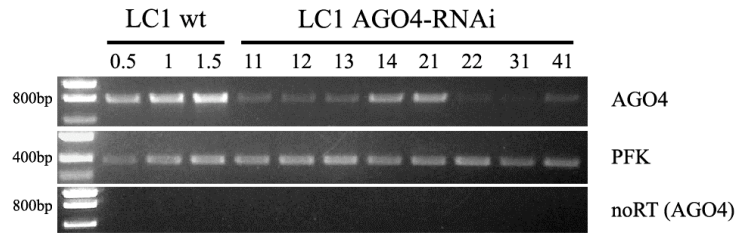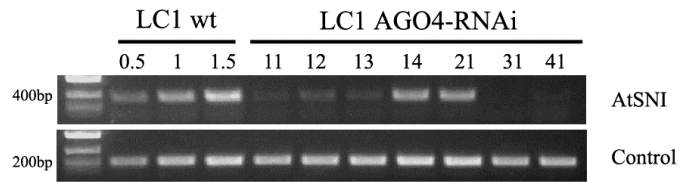
observations in the *A. thaliana ago4-1* mutant line (Chapter IV), individual *31* also displays the most severe effect in the ac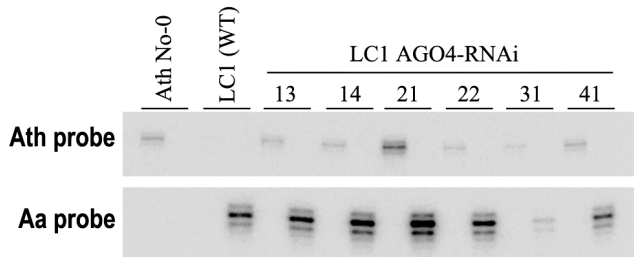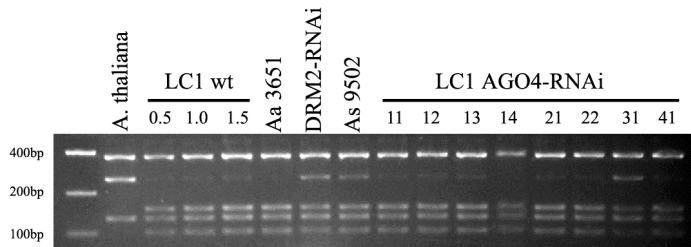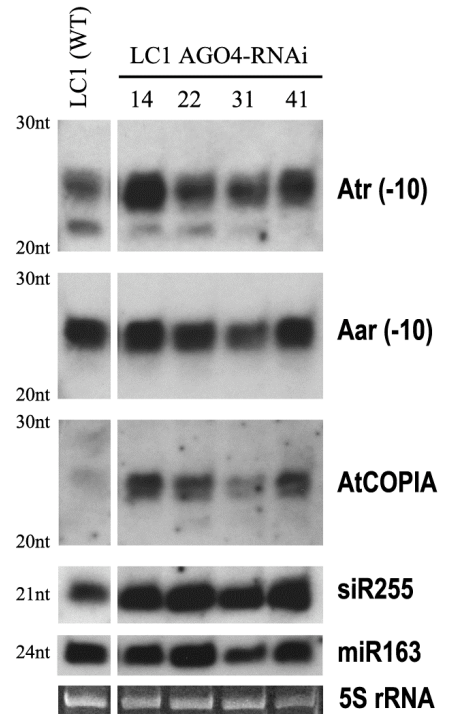cumulation of *AtCOPIA*-derived siRNAs amongst the analyzed transformants (fig. 5.6e). Due to lack of hybridization signal in LC1 WT it is impossible to evaluate, at this point, to what extent knock-down of *AGO4* compromises *AtCOPIA* siRNA biogenesis.

V.3. Interphase heterochromatin organization in A. suecica is modulated by RNAi pathway genes.

In *A. thaliana*, null alleles of *nrpd1a* and *nrpd2a* display altered H3$^{\text{dimethyl}}$K9 patterns in interphase nuclei (Onodera *et al.*, 2005). In order to assess if heterochromatin organization is modulated by the endogenous heterochromatic RNAi pathway in A. suecica, H3$^{\text{dimethyl}}$K9 interphase organization was determined in the progeny of selected T$_1$ transformants by *in situ* immunolocalization (table 5.1; fig. 5.7). T$_1$ transformants which showed the best correlation between target mRNA knock-down, *AtSN1 loci* demethylation and deleterious effects on smRNA accumulation were selected for the analysis. The H3$^{\text{dimethyl}}$K9 interphase organization observed in WT LC1 nuclei agree with previously published results (Earley *et al.*, 2006; Lawrence *et al.*, 2004). In the majority of analyzed nuclei (84.4%) H3$^{\text{dimethyl}}$K9 overlaps chromocenters organized in condensed and discrete *foci*. This interphase pattern was also observed in AGO4-RNAi nuclei indicating that AGO4 does not play a significant role in heterochromatin organization. In contrast, RDR2-, NRPD2- and HEN1-RNAi lines display significant alterations in H3$^{\text{dimethyl}}$K9 interphase organization patterns (table 5.1; fig 5.7). Typically, the WT discrete and condensed immunolocalization signal is disrupted into a decondensed and fuzzy pattern. Importantly, the H3$^{\text{dimethyl}}$K9 signal is no longer restricted to the DAPI positive chromocenters but extends into the nucleoplasm in these transgenic lines.

Table 5.1 - Frequency (%) of nuclei observed with distinct H3$^{\text{dimethyl}}$K9 interphase organization patterns in *A. suecica* LC1 (LC1) and progeny of T$_1$ transformants for *RDR2-RNAi* (*43*), *NRPD2a-RNAi* (*41*), *AGO4-RNAi* (*31*) and *HEN1-RNAi* (*14*) lines. H3$^{\text{dimethyl}}$K9 WT pattern is defined according to Lawrence *et al.* (2004). χ^2-test was used to compare distributions between transgenic lines and WT (d.f.=2). Null hypothesis was not rejected if p>0.5.

		Genotype				
		LC1	*NRPD2a*	*RDR2*	*AGO4*	*HEN1*
H3m2K9	**WT**	84.4	8.8	52.0	76.8	47.1
pattern	**Disturbed**	15.6	91.2	48.0	23.2	52.9
	χ^2		218.4 (p≤0.001)	50.1 (p≤0.001)	3.4 (p≤0.1)	63.6 (p≤0.001)
	n	173	205	304	198	295

Figure 5.7 – Immunolocalization of histone H3 dimethylated on lysine 9 (green) in leaf interphase nuclei of *A. suecica* LC1 and NRPD2-, RDR2-, AGO4- and HEN1-RNAi lines. Images shown are representative of the H3$^{\text{dimethyl}}$K9 pattern more frequently observed in each of the analyzed lines (see also table V.1 for quantitative data). DNA was counterstained with DAPI. (size bar = 5μm).
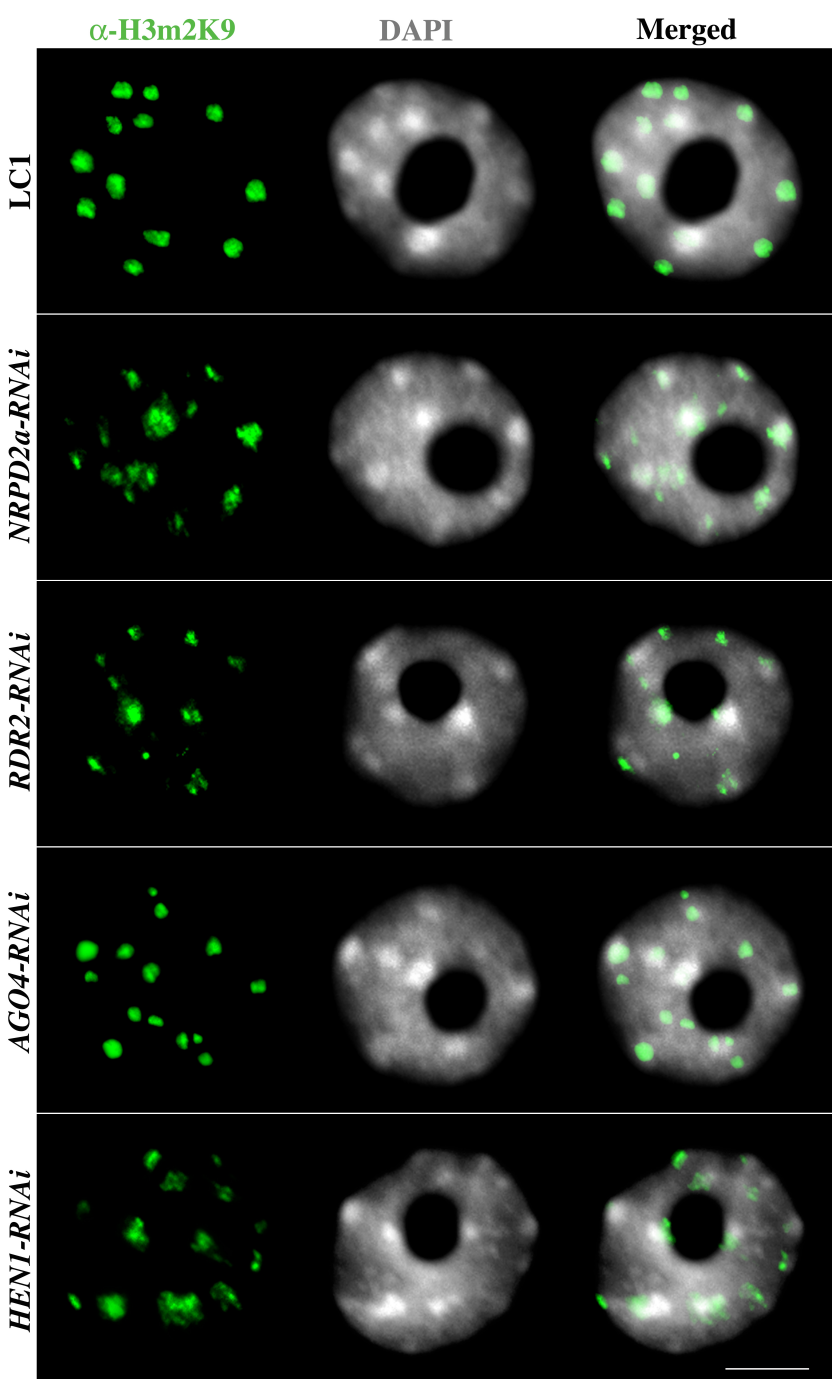
Reactivation of *A. thaliana* 45S rRNA genes in *A. suecica* has been linked to *locus* decondensation and loss of association with the H3[dimethyl]K9 heterochromatic mark (Earley *et al.*, 2006; Lawrence *et al.*, 2004). It was previously observed that, in an *A. thaliana* genomic background, NOR chromatin is decondensed in *rdr2-1* and *nrpd2a/nrpd2b* mutant backgrounds (Chapter IV). RNAi mediated knock-down of RDR2 and NRPD2 mRNAs in *A. suecica* LC1 results in decondensation of *A. thaliana* NORs, shifting from highly condensed *foci* observed in WT nuclei to a disrupted chromatin organization where more than three rRNA foci are frequently identified (>58.6%) (table 5.2; fig. 5.8). Importantly, NOR *loci* partially loose association with H3[dimethyl]K9 in the majority of interphase nuclei analyzed by double immunolocalization; DNA:DNA *in situ* hybridization.

Table 5.2 - Frequency (%) of nuclei observed with distinct numbers of *A. thaliana* rRNA FISH signals in *A. suecica* LC1 (LC1) and progeny of T_1 transformants of *A. suecica RDR2-RNAi* (*43*) and *A. suecica NRPD2-RNAi* (*41*) lines. *A. thaliana* NORs are considered partially colocalized with H3[dimethyl]K9 when decondensed (≥ 3 *foci*) and colocalized when condensed (≤ 2 *foci*). χ^2-test was used to compare distributions between transgenic lines and WT (d.f.=2). Null hypothesis was not rejected if p>0.5.

	45S rDNA *foci* number (%)			χ^2	*n*
	<2	=2	≥ 3		
LC1 (wt)	18.3	67.1	14.6		82
NRPD2-RNAi	10.8	24.6	64.6	42.3 (p≤0.001)	175
RDR2-RNAi	14.0	27.4	58.6	57.4 (p≤0.001)	128
	total		*partial*		

H3m2K9 vs 45S rDNA *foci* colocalization

Figure 5.8 – Double immunolocalization and DNA FISH in leaf interphase nuclei of *A. suecica* LC1 (LC1) NRPD2a- and RDR2-RNAi. Images reflect the most frequently observed organization patterns observed for H3$^{\text{dimethyl}}$K9 (green) and *A. thaliana* derived 45S rDNA *loci* (red) (see also table V.2 for quantitative data). 45S rDNA was hybridized with a biotin labeled *A. thaliana* specific 45S rDNA IGS probe (pAt.2). In merged images yellow signals correspond to regions where *A. thaliana* 45S rDNA chromatin is associated with H3$^{\text{dimethyl}}$K9. DNA was counterstained with DAPI. (size bar = 5μm).
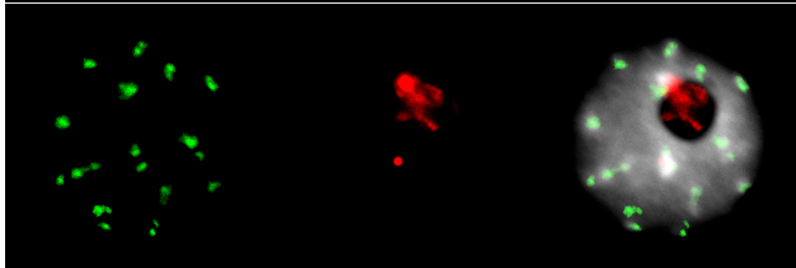
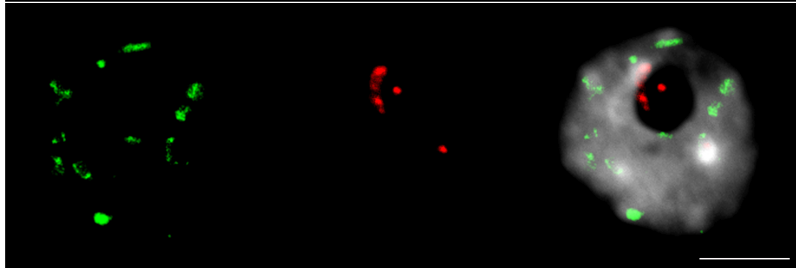| α-H3m2K9	AtNORs	Merged+DAPI

LC1

NRPD2a-RNAi

RDR2-RNAi

V.4. Discussion.

In *A. suecica*, nucleolar dominance is linked to the epigenetic silencing of *A. thaliana* derived 45S rRNA genes. Erasure, or impairment, of DNA methylation or histone deacetylation in the 45S rRNA gene promoter region correlates with disruption of nucleolar dominance (Chen *et al.*, 1997a), a phenotype also observed upon RNAi mediated knock-down of the DNA methyltransferase *DRM2* (Preuss and Pikaard, personal communication) or the histone deacetylases *HDA6* or *HDT1* (Earley *et al.*, 2006; Lawrence *et al.*, 2004). Importantly MET1- and DDM1-RNAi knock-down lines do not display reactivation of *A. thaliana* 45S rRNA genes indicating that DNA methylation of cytosines in symmetric sequence contexts is not relevant for the epigenetic control of rRNA genes (Lawrence, Preuss and Pikaard, personal communication). Furthermore, nucleolar dominance is developmentally regulated and occurs in leaf tissue in *Arabidopsis* and *Brassica sp.* but not in floral tissues or cotyledons (Chen and Pikaard, 1997; Lawrence, Pontes, Viegas and Pikaard, personal communication). 45S rRNA gene activity is associated with the metabolic needs of the cell and, whereas rapidly dividing undifferentiated cells of fast growing shoots or meristems require higher transcription rates of rRNA for ribosome assembly, in differentiated cells the demand for 45S rRNA precursors can be lower. Differentiation leads to a reduction of the number of rRNA genes that are in an open chromatin state and therefore able to support transcription by pol I, in a process possibly also modulated by the available pool (Li *et al.*, 2006) and gene occupancy density of pol I at a given gene unit (Zilberman *et al.*, 2007). This mechanism of rRNA gene control renders irrelevant 45S rRNA copy number in a diploid genome or in an autopolyploid context (Muscarella *et al.*, 1985; Rivin *et al.*, 1986). Together these observations suggest that a subset of 45S rRNA genes is epigenetically controlled to be in an "on" or "off" state which, in *A. suecica*, requires the activity of DRM2 and histone deacetylases HDT1 and HDA6.

Nucleolar dominance is a phenomenon observed in interspecific hybrids, including allopolyploids. In plant hybrids nucleolar dominance is epigenetically regulated by an unknown choice mechanism which selects one parental 45S rRNA gene subset for silencing (Chen *et al.*, 1997a; Chen *et al.*, 1998; Earley *et al.*, 2006; Lawrence *et al.*, 2004; Neves *et al.*, 1995; Vieira *et al.*, 1990). In *A. suecica*, nucleolar dominance is intimately associated with genome evolution. Newly formed synthetic allopolyploids display variability between the expression balance of *A. thaliana* and *A. arenosa* 45S rRNA genes, ranging from bidirectional nucleolar dominance to codominance (Chen *et al.*, 1998). These observations can be partially explained by an also variable number of parental NOR *loci* in the evolving hybrid genome (Pontes *et al.*, 2004). Backcross of a natural *A. suecica* line to the diploid *A. thaliana* also results in nucleolar dominance disruption (Chen *et al.*, 1998). Together with the fact that

A. thaliana / A. arenosa NOR balance in the natural *A. suecica* tetraploid hybrid genome is 2/6 (Pontes *et al.*, 2003), this indicates that parental gene dosage influences nucleolar dominance establishment in *Arabidopsis*. It is nevertheless insufficient, as two *A. suecica* natural accessions with high degree of genomic similarity and the same parental NOR *loci* balance display distinct NOR transcription regulation: *A. suecica* LC1 displays nucleolar dominance whereas *A. thaliana* derived 45S rRNA transcripts are readily detected in *A. suecica* 9502 (Pontes *et al.*, 2003).

DNA sequence polymorphisms in the upstream region relative to the 45S rRNA transcription start site between *A. arenosa* and *A. thaliana* rDNA *loci* allow the distinction of parental specific smRNA species. The evaluation of siRNA accumulation levels in *A. suecica* lines displaying nucleolar dominance disruption demonstrates that the endogenous heterochromatic siRNA pathway is not the sole mechanism involved in the establishment of this epigenetic phenomenon. This conclusion arises from the observation that HDT1-RNAi lines display WT 24nt siRNA accumulation levels. In contrast, HDA6-RNAi shows a dramatic increase in rRNA-homologous 21 and 24nt siRNAs, as is also observed in *A. thaliana axe1-5* (*hda6*) mutants. This histone deacetylase has been linked to endogenous RNA mediated gene silencing pathways (Aufsatz *et al.*, 2002; Murfett *et al.*, 2001) as is required for rDNA *locus* chromatin organization and DNA methylation in both *A. thaliana* and *A. suecica* (Earley *et al.*, 2006; Probst *et al.*, 2004). HDA6 is apparently able to deacetylate multiple (and possibly all) acetylated lysines in H3 and H4 tails *in vitro* as well as *in vivo* (Earley et al., 2006) which points to a potentially broad spectrum of chromatin related mechanisms that may or may not be related to RNAi. Histone acetylation has been linked to active transcription (Workman and Kingston, 1998) and to potentiate loosening of inter- or intra-nucleossomal DNA-histone interactions making acetylated histones easier to displace from DNA (Reinke and Horz, 2003; Zhao *et al.*, 2005). Deacetylation is required to maintain chromatin in a stable conformation which in turn is repressive for RNA polymerase initiation complex formation and transcription. In HDA6-RNAi lines, histone hyperacetylation occurs, which could allow cryptic promoter-like sequences to function as transcription start sites, as has been documented in yeast (Earley *et al.*, 2006; Carrozza *et al.*, 2005; Joshi and Struhl, 2005). Hence, the increase in accumulation levels of siRNAs in *axe1-5* and HDA6-RNAi plants could be the result of an overall increase in the transcription rate of non-coding RNAs which, in turn, feed into endogenous RNAi pathways. In the context of nucleolar dominance, failure to deacetylate histones in the A. thaliana 45S rRNA gene promoter region would lead to the inability of repressive complexes to silence this parental subset of rRNA genes, thereby translating into nucleolar dominance disruption, likely independently or upstream of RNAi pathway mechanisms.

The endogenous heterochromatic siRNA pathway is apparently also implicated in nucleolar dominance. *DRM2* is a member of the pathway required for 45S rRNA core promoter-homologous

24nt siRNAs in *A. thaliana* (Chapter IV; Pontes et al. 2006). Knock-down of this DNA methyltransferase in *A. suecica* disrupts nucleolar dominance which correlates with reduced accumulation of both *A. thaliana* and *A. arenosa* homologous siRNAs. Importantly, the same smRNA phenotype is also observed in the natural accession 9502 which also displays a strong reduction in siR1003 (5S rRNA) accumulation levels. Disruption of nucleolar dominance in this natural accession has been previously associated with *A. thaliana* NOR decondensation during interphase (Pontes *et al.*, 2003), a nuclear phenotype intimately associated with *A. thaliana* 45S rRNA gene reactivation in *A. suecica* (Earley *et al.*, 2006; Lawrence *et al.*, 2004; Pontes *et al.*, 2003). Furthermore, altered nuclear chromatin organization is observed as a consequence of disrupting the 24nt siRNA pathway in *A. thaliana* (Onodera *et al.*, 2005), in which NOR decondensation also occurs (Chapter IV). It is therefore feasible to consider that the heterochromatic RNAi pathway is involved in the establishment or maintenance of nucleolar dominance.

According to the S1 nuclease protection assay results, only a subset of 45S rRNA genes is apparently under the control of RNAi silencing mechanisms. Nevertheless, it substantiates the challenging hypothesis that the heterochromatic RNAi pathway is part of the choice mechanism which dictates the parental set targeted for silencing in nucleolar dominance. Unfortunately, the effort to disrupt the 24nt endogenous RNAi pathway in *A. suecica*, by RNAi induced knock-down of identified players involved in 45S rRNA siRNA biogenesis (Chapter IV), was insufficient to definitely correlate this pathway with the establishment of nucleolar dominance. Analysis of more transgenic RNAi lines, or knock-down mediated using artificial miRNAs rather than double-stranded hairpin RNA triggers, will be necessary to better evaluate the role of the heterochromatic siRNA pathway in nucleolar dominance. Nevertheless, the observation that for all analyzed knock-down targets, higher levels of *A. thaliana* 45S precursor rRNAs were always identified in transformants which combine good knock-down efficiency with DNA demethylation of *AtSNI loci* and reduction of 24nt siRNA accumulation is encouraging. Furthermore, cytological analysis of interphase nuclei also demonstrated that these same transformants affect H3$^{\text{dimethyl}}$K9 organization patterns and, importantly, that *A. thaliana* NORs partially lose association to this heterochromatic mark in RDR2- and NRPD2-RNAi transformants. The same interphase phenotype is also observed in HDA6-RNAi linking it to nucleolar dominance disruption (Earley *et al.*, 2006).

It is possible that the variability of the effects observed within each RNAi knock-down line generated for this study is the result of the inherent particularities of the 24nt siRNA pathway. This multi-step pathway is self-reinforcing which indicates that multiple rounds of RNA precursor transcription and processing are required to efficiently establish the heterochromatic state at target genome regions. Hence it is feasible that knock-down efficiency, which does not equate the effect of

null alleles, mainly delays the effectiveness of the silencing pathway as a whole. In such a scenario, the replication of the experiment in *A. suecica* 9502 instead of an LC1 genomic background could provide more conclusive evidence of the involvement of the siRNA heterochromatic pathway in nucleolar dominance. The reduced levels of rRNA homologous siRNAs in 9502 indicate that its biogenesis is compromised. In light of the analysis performed in *A. thaliana* (Chapter IV) showing that *ago4-1* displays a mild effect on siRNA species homologous to 5S and 45S rRNA genes, it is apparent that an amplification step occurs before siRNA loading into AGO4, possibly involving another, and yet to be determined, Argonaute family member. This hypothesis would not be unique in the context of our present knowledge of endogenous RNAi pathways, as a requirement for at least two distinct Argonaute proteins for the biogenesis of specific smRNA species have been reported in *C. elegans* and *A. thaliana* (Adenot *et al.*, 2006; Gasciolli *et al.*, 2005; Qi *et al.*, 2006). AGO4 is also known to play two distinct functions related to its ability to bind 23-24nt smRNAs (Li *et al.*, 2006; Qi *et al.*, 2006); in some cases acting as an RNA slicer in addition to targeting DNA modifications and silencing of target *loci* (Yigit *et al.*, 2006). Therefore, replication of the RNAi induced knock-down of 24nt siRNA heterochromatic pathway genes in *A. suecica* 9502 could result in further impairment of the pathway, as two steps would be affected. The additive effect obtained would be expected to severely compromise siRNA biogenesis and produce a stronger correlation between 45S rRNA siRNAs and nucleolar dominance, if such hypothesis is proven correct.

Could siRNA mediated gene silencing be required for parent-specific silencing of NORs in the context of nucleolar dominance in *A. suecica*? One interesting difference in the IGS regions of *A. thaliana* and *A. arenosa* 45S rRNA genes is that *A. thaliana* possesses three putative pol I promoter regions for each gene unit (two spacer promoters in addition to the gene promoter) whereas *A. arenosa* only has one, the 45S rRNA core promoter. This characteristic of the IGS regions could potentiate a higher tendency of *A. thaliana* rRNA genes to be subjected to heterochromatinization by a siRNA based targeting mechanism if spacer promoters provide RNA transcripts that give rise to siRNAs targeting regulatory elements in the spacer. Alternatively, other sequence differences may lead to higher densities of heterochromatin like regions in the *A. thaliana* 45S rRNA IGS, when compared to the single restricted core promoter region in *A. arenosa* based on a siRNA-mediated mechanism. The downstream effect in terms of nucleolar dominance establishment would be the result of the higher capability of *A. arenosa* derived rRNA genes to compete for the available transcription machinery.

VI – Conclusions and future prospects.

The heterochromatic siRNA pathway is required to establish heterochromatic marks in a mechanism involving DNA methylation and histone post-translational modifications. Central to this epigenetic silencing pathway are the 24nt siRNA molecules which guide silencing complexes in a sequence specific manner to DNA target regions. Biogenesis of these small non-coding RNAs is exclusively dependent on processing of RNA precursors by RDR2 and DCL3 (Xie *et al.*, 2004) in a step that distinguishes this pathway from all other endogenous RNAi pathways that are know to occur in a plant cell. Both pol IVa and pol IVb, AGO4 and DRD1 are also required for the proper functioning of the 24nt siRNA pathway, with pol IVa primarily responsible for the generation of RNA precursors and the other activities participating in targeting and establishment of epigenetic marks in the chromatin (Herr *et al.*, 2005; Kanno *et al.*, 2004; 2005; Onodera *et al.*, 2005; Pontes *et al.*, 2006; Zilberman *et al.*, 2003; 2004).

In *Arabidopsis sp*, 45S rRNA genes are under epigenetic regulation, which in the diploid *A. thaliana* functions as a dosage control mechanism and in the tetraploid *A. suecica* can lead to the establishment of nucleolar dominance (Lawrence *et al.*, 2004). Characterization of smRNAs homologous to 45S rRNA genes allowed the identification of a discrete interval of preferential accumulation of 24nt siRNA species overlapping the core promoter region. Biogenesis of these smRNA species is dependent on the activity of all the genes implicated in the 24nt siRNA heterochromatic pathway. Core promoter siRNAs associate with AGO4 (Pontes *et al.*, 2006) and siRNA accumulation is strongly impaired in *drm2* mutants indicating that the potential for siRNA-directed DNA methylation of 45S rRNA promoter regions exists. The obtained results regarding nuclear interphase organization of 45S rRNA *loci* demonstrate that the endogenous 24nt siRNA pathway participates in higher order chromatin organization, as also observed for 5S rRNA *loci* (Onodera *et al.*, 2005), but no correlation with DNA methylation levels or changes in 45S rRNA transcript levels were obtained in this study. CHIP-based (**Ch**romatin **I**mmuno**p**recipitation) assays may provide the missing piece of the puzzle based on our current understanding that the number of pol I complexes engaged in transcription as well as the number of genes in the "on" state dictates overall pre-rRNA transcript levels. It has been shown that chemically induced DNA hypomethylation and histone deacetylation increase the set of 45S rRNA genes in a transcriptionally competent chromatin conformation (Earley *et al.*, 2006; Lawrence *et al.*, 2004) and that pol I preferentially associates with DNA-hypomethylated and H3$^{\text{trimethyl}}$K4 enriched promoters (Lawrence *et al.*, 2004). One could expect that the pol I pool could be redistributed by a larger set of rRNA genes as a result of knock-out of the RNAi pathway in *A. thaliana* and that chromatin pull-downs targeting pol I complexes would demonstrate a decrease in the density of 45S rRNA promoter occupancy while an enrichment in the

euchromatic H3^{trimethyl}K4 mark would be expected to occur. DNA methylation levels on 45S rRNA gene promoters could also be assessed by use of CHIP-CHOP PCR (Lawrence *et al.*, 2004) and directly correlated with transcriptional status by combining the analysis with immunoprecipitation of chromatin using an anti-pol I antibody.

Nucleosome remodeling has also been linked to the modulation of rRNA gene expression. NoRC - mediated nucleosome remodeling activity in mammals was shown to re-position nucleosomes over the rRNA gene transcription start site. Active genes have nucleotides from -157 to -2 wrapped around the promoter bound nucleosome. NoRC activity slides the nucleosome 25 nucleotides downstream, resulting in rRNA silencing (Li *et al.*, 2006). It would be interesting to determine nucleosome positioning in active *vs* inactive rRNA genes in *A. thaliana* and to see if any correlation exists between the 24nt siRNA heterochromatic pathway (which includes the putative chromatin remodeller DRD1), the discrete distribution of 45S rRNA core promoter homologous siRNAs, and nucleosome positioning.

In order to understand the role of 24nt siRNAs in modulating 45S rRNA gene expression in the context of nucleolar dominance, RNAi mediated knock-down of the endogenous RNA dependent heterochromatic pathway was performed in the tetraploid *A. suecica* LC1, in which nucleolar dominance occurs. The analysis of RNAi lines known to display nucleolar dominance disruption (Earley *et al.*, 2006; Lawrence *et al.*, 2004; Preuss and Pikaard, unpublished) and the *A. suecica* natural accession 9502 (Pontes *et al.*, 2003) demonstrated that, although not its sole determinant, 24nt siRNAs are likely involved in the modulation of this epigenetic phenomenon. Both 9502 and DRM2-RNAi display reduced levels of 45S rRNA homologous 24nt siRNAs, a phenotype that was also observed in T$_1$ transformants where RNAi induced knock-down of *NRPD2a*, *RDR2*, *DCL3* and *HEN1* mRNA was most efficient These lines displayed nucleolar dominance disruption, *AtSN1* DNA demethylation, interphase decondensation and, in the case of NRPD2- and RDR2-RNAi individuals, loss of complete association of the H3^{dimethyl}K9 heterochromatic mark with *A. thaliana* 45S rRNA *loci*. This latter interphase organization phenotype was previously associated with *A. thaliana* 45S rRNA gene transcriptional activity (Earley *et al.*, 2006). Nevertheless, the great variability encountered between T$_1$ transformants indicates that further analysis is required to definitely implicate the 24nt siRNA heterochromatic pathway in nucleolar dominance establishment. Replication of the experiment with recently available tools which rely on the miRNA machinery to efficiently induce target mRNA decay (Schwab *et al.*, 2006) could provide the further insights required.

The presented results also underline the possible requirement of other Argonaute proteins for the biogenesis of 45S rRNA homologous 24nt siRNAs. In *A. thaliana*, *ago4-1* displays near WT siRNA

accumulation levels and such is also the case for 5S rRNA-derived siRNAs (note that AGO4 is required for DNA methylation of 5S rRNA genes), a phenotype not observed in loss of function mutants of other pathway members or for other *loci* where AGO4 is determinant in siRNA biogenesis (e.g. *AtSN1*). Further characterization of other Argonaute proteins is required and, apart from resorting to loss of function mutants in an *A. thaliana* background, *A. suecica* 9502 can also prove to be a useful tool in this analysis. Both 5S and 45S rRNA-derived siRNA accumulation is strongly compromised in this line and evaluation of the expression levels and characterization of *Argonaute* genes in an *A. suecica* genomic background could help determine which of the other known nine Argonaute family members is required for rRNA siRNA biogenesis and accumulation.

VII – Conclusões e perspectivas futuras.

A via endógena de siRNAs heterocromáticos participa na modulação da heterocromatina pelo estabelecimento de marcas de metilação no DNA e modificações pós-tradução de histonas. Os siRNAs de 24nt constituem um peça fundamental nesta via de silenciamento epigenético dirigido por serem responsáveis pelo encaminhamento dos complexos silenciadores a regiões de DNA de sequência nucleotídica homóloga à sua. A biogénese destas moléculas de RNA não-codificante depende exclusivamente do processamento de percursores de RNA por parte das proteinas RDR2 e DCL3 (Xie *et al.*, 2004), num passo que distingue esta via heterocromática de todos os outros mecanismos de RNA de interferência que ocorrem numa célula vegetal. Entre outros membros identificados incluem-se os complexos pol IVa e pol IVb e as proteinas AGO4 e DRD1, sendo o complexo pol IVa responsável pela síntese da molécula de RNA precursora, enquanto os outros membros participam no estabelecimento de marcas epigenéticas ao nível da cromatina alvo (Herr *et al.*, 2005; Kanno *et al.*, 2004; 2005; Onodera *et al.*, 2005; Pontes *et al.*, 2006; Zilberman *et al.*, 2003; 2004).

Em *Arabidopsis sp.*, a expressão dos genes ribossomais (45S) encontra-se sobre regulação epigenética, funcionando como mecanismo de controlo de dosagem no diplóide *A. thaliana* e sendo responsável pelo estabelecimento da dominância nucleolar no tetraplóide *A. suecica* (Lawrence *et al.*, 2004). A caracterização de moléculas de smRNAs homólogas aos genes 45S permitiu a identificação de uma região discreta de acumulação preferencial de siRNAs de 24nt correspondente à região do promotor dos genes de rDNA. Estes smRNAs associam-se com AGO4 (Pontes *et al.* 2006) e a sua biogénese é fortemente comprometida em linhas mutantes *drm2*, indiciando que os siRNAs homólogos à região do promotor serão potenciais moduladores do controlo epigenético dos genes ribossomais. Os resultados obtidos demonstram que a via de siRNAs heterocromáticos participa na conformação da cromatina ribossomal, como anteriormente documentado relativamente aos genes ribossomais 5S (Onodera *et al.*, 2005). No entanto, as alterações na organização interfásica da cromatina não apresentam uma correlação directa com alterações nos níveis de metilação de DNA na região do promotor nem com alterações a nível global de acumulação de transcritos provenientes dos genes 45S. O recurso a técnicas de immuno-precepitação de cromatina (CHIP) poderá fornecer a prova defenitiva da importância das moléculas de siRNAs na modulação epigenética dos genes ribossomais. O nível de expressão endógeno dos genes ribossomais 45S resulta de um balanço entre o número de genes transcripcionalmente competentes e a "pool" de pol I disponivel para a sua transcrição. O uso de agentes químicos que induzem hipometilação de DNA ou hiperacetilação de histonas resulta no aumento do número de genes ribossomais associados a marcas eucromáticas (Earley *et al.*, 2006; Lawrence *et al.*, 2004), fracção a que se associam os complexos de pol I (Lawrence *et al.*, 2004). O bloqueamento da via de siRNAs heterocromáticos, embora aumente a

fracção de genes ribossomais eucromáticos, poderá resultar na redistribuição de complexos pol I sem afectar necessariamente os níveis totais de transcrição do pré-rRNA. Sendo esta interpretação correcta, por immunoprecipitação da cromatina ribossomal via pol I, deverá ser possível observar um decréscimo na densidade de ocupação desta RNA polimerase nos promotores dos genes 45S e um enriquecimento no nível de associação destas regiões a marcas eucromáticas (via imuno-precipitação com um anticorpo anti-H3^{trimetilada}K4). Alterações nos níveis de metilação da região do promotor, e sua correlação com marcas epigenéticas ou associação a pol I, poderão ser também avaliadas pela combinação de imuno-precipitação com digestão enzimática de DNA metilado (CHIP-CHOP PCR; Lawrence *et al.*, 2004).

O posicionamento dos nucleossomas está também associado à regulação da expressão génica dos genes ribossomais. A actividade remodeladora do complexo repressivo NoRC em mamíferos reposiciona nucleossomas por forma a que estes englobem o local de início de transcrição do rRNA percursor (45S). Nos genes ribossomais activos, o nucleossoma engloba a sequência de DNA entre as posições -152 a -2. A actividade do NoRC desloca o nucleossoma 25bp a jusante, bloqueando a acessibilidade de factores de transcrição / pol I ao promotor, resultando no silenciamento do gene ribossomal (Li *et al.*, 2006). Seria interessante determinar o posicionamento dos nucleossomas na região promotora de genes ribossomais activos e inactivos por forma a averiguar se existe alguma correlação entre este posicionamente, a via de siRNAs heterocromáticos (da qual faz parte a proteína remodeladora de cromatina DRD1) e a distribuição de siRNAs homólogos ao promotor dos genes de rDNA 45S.

Por forma a estudar o papel dos siRNAs (24nt) na modulação da expressão génica dos genes ribossomais no contexto da dominância nucleolar, utilizou-se a tecnologia de RNA de interferência para induzir o "knock-down" de genes pertencentes à via de siRNAs heterocromáticos no híbrido *A. suecica*. A análise de linhas transgénicas HDT1-, HDA6-, DRM2-RNAi (Earley *et al.*, 2006; Lawrence *et al.*, 2004; Preuss and Pikaard, comunicação) e do tetraplóide natural *A. suecica* 9502, em que não se observa dominância nucleolar, demonstrou que a via de siRNAs heterocromáticos está também envolvida no estabelecimento deste fenómeno epigenético, embora não sendo o seu único factor determinante. *A. suecica* 9502 e DRM2-RNAi apresentam níveis de acumulação reduzidos de siRNAs homólogos às regiões promotoras dos genes ribossomais. O mesmo fenótipo de reduzida acumulação de siRNAs e do não estabelecimento de dominância nucleolar é também observado em plantas transgénicas NRPD2-, RDR2-, DCL3- e HEN1-RNAi onde, a um eficiente "knock-down" do mRNA do gene alvo, se associa a redução dos níveis de metilação nos *loci AtSNI* e descondensação dos NORs de *A. thaliana* em interfase. No caso de transgénicos NRPD2- e RDR2-RNAi, a

descondensação da cromatina ribossomal em interfase resulta numa perda de associação de genes de rDNA a H3$^{\text{dimetilada}}$K9, um fenótipo interfásico associado à reactivação de genes ribossomais de *A. thaliana* em *A. suecica* (Earley *et al.*, 2006). No entanto, e devido à variabilidade encontrada entre transformantes de um mesmo "knock-down" (ex. RDR2-RNAi), mais estudos serão necessários para correlacionar a via de siRNAs heterocromáticos com o estabelecimento da dominância nucleolar em *A. suecica*. A repetição da abordagem realizada recorrendo a novas ferramentas de RNAi, que dependem da via endógena de miRNAs (Schwab *et al.*, 2006) para induzir o "knock down" dos genes alvo, poderá proporcionar a prova defenitiva.

Os resultados obtidos evidenciam que outra proteína da família Argonauta poderá estar envolvida na biogénese de siRNAs homólogos a genes ribossomais. Em *A. thaliana*, mutações no gene *AGO4* (*ago4-1*) não resultam num decréscimo assinalável dos níveis de acumulação de siRNAs de 24nt derivados dos genes de rDNA 45S e 5S, embora para o último se verifique o consequente decréscimo nos níveis de metilação do DNA. Este fenótipo não é observado em linhas de *A. thaliana* com alelos nulos de outros membros da via de siRNAs (24nt) ou no caso de siRNAs fortemente dependentes da função da AGO4 (ex. *AtSN1*). A caracterização do papel de outras Argonautas na biogénese de siRNAs de genes ribossomais é portanto necessária de forma a identificar que outro membro desta família poderá estar envolvido. Esta análise poderá não ser restringida aos recursos genéticos disponíveis no background genómico de *A. thaliana*, englobando também *A. suecica* 9502. Este híbrido apresenta uma forte redução nos níveis de acumulação de siRNA homólogos a genes de rDNA (5S e 45S) e a avaliação dos níveis de expressão e/ou caracterização de genes *AGO* neste contexto genómico poderá ajudar a determinar qual dos outros nove membros desta familia é necessário para a biogénese desta classe de siRNAs.

VIII – References.

Adenot, X., Elmayan, T., Lauressergues, D., Boutet, S., Bouche, N., Gasciolli, V. and Vaucheret, H. (2006). "DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7." *Curr Biol* **16**(9): 927-932.

Agius, F., Kapoor, A. and Zhu, J. K. (2006). "Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation." *Proc Natl Acad Sci U S A* **103**(31): 11796-11801.

Ahmad, K. and Henikoff, S. (2002). "The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly." *Mol Cell* **9**(6): 1191-1200.

Albert, A. C., Denton, M., Kermekchiev, M. and Pikaard, C. S. (1999). "Histone acetyltransferase and protein kinase activities copurify with a putative Xenopus RNA polymerase I holoenzyme self-sufficient for promoter-dependent transcription." *Mol Cell Biol* **19**(1): 796-806.

Allen, E., Xie, Z., Gustafson, A. M. and Carrington, J. C. (2005). "microRNA-directed phasing during trans-acting siRNA biogenesis in plants." *Cell* **121**(2): 207-221.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C. and Ecker, J. R. (2003). "Genome-wide insertional mutagenesis of Arabidopsis thaliana." *Science* **301**(5633): 653-657.

Amarzguioui, M., Holen, T., Babaie, E. and Prydz, H. (2003). "Tolerance for mutations and chemical modifications in a siRNA." *Nucleic Acids Res* **31**(2): 589-595.

Ambros, V. (2004). "The functions of animal microRNAs." *Nature* **431**(7006): 350-355.

Arabidopsis-Genome-Initiative (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana." *Nature* **408**(6814): 796-815.

Aufsatz, W., Mette, M. F., Matzke, A. J. and Matzke, M. (2004). "The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides." *Plant Mol Biol* **54**(6): 793-804.

Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, M. and Matzke, A. J. (2002). "HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA." *Embo J* **21**(24): 6832-6841.

Axtell, M. J., Jan, C., Rajagopalan, R. and Bartel, D. P. (2006). "A two-hit trigger for siRNA biogenesis in plants." *Cell* **127**(3): 565-577.

Bannister, A. J., Schneider, R. and Kouzarides, T. (2002). "Histone methylation: dynamic or static?" *Cell* **109**(7): 801-806.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. and Kouzarides, T. (2001). "Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain." *Nature* **410**(6824): 120-124.

Bartee, L., Malagnac, F. and Bender, J. (2001). "Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene." *Genes Dev* **15**(14): 1753-1758.

Bartel, D.P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." *Cell* **116**:281-297.

Baulcombe D. (20040. "RNA silencing in plants." *Nature* **431**:356-363.

Baumberger, N. and Baulcombe, D. C. (2005). "Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs." *Proc Natl Acad Sci U S A* **102**(33): 11928-11933.

Belmont, A. S., Dietzel, S., Nye, A. C., Strukov, Y. G. and Tumbar, T. (1999). "Large-scale chromatin structure and function." *Curr Opin Cell Biol* **11**(3): 307-311.

Bennetzen, J. L. (1998). "The structure and evolution of angiosperm nuclear genomes." *Curr Opin Plant Biol* **1**(2): 103-108.

Bennetzen, J.L. (1996). "The contributions of retroelements to plant genome organization, function and evolution." *Trends Microbiol.* **4**(9):347-53.

Bernard, P., Maure, J. F., Partridge, J. F., Genier, S., Javerzat, J. P. and Allshire, R. C. (2001). "Requirement of heterochromatin for cohesion at centromeres." *Science* **294**(5551): 2539-2542.

Bernstein, E. and Allis, C. D. (2005). "RNA meets chromatin." *Genes Dev* **19**(14): 1635-1655.

Bernstein, E., Denli, A. M. and Hannon, G. J. (2001). "The rest is silence." *Rna* **7**(11): 1509-1521.

Birve, A., Sengupta, A. K., Beuchle, D., Larsson, J., Kennison, J. A., Rasmuson-Lestander, A. and Muller, J. (2001). "Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants." *Development* **128**(17): 3371-3379.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M. and Benning, C. (1998). "AGO1 defines a novel locus of Arabidopsis controlling leaf development." *Embo J* **17**(1): 170-180.

Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R. and Zhu, J. K. (2005). "Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis." *Cell* **123**(7): 1279-1291.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., Roth, R., George, D., Eletr, S., Albrecht, G., Vermaas, E., Williams, S. R., Moon, K., Burcham, T., Pallas, M., DuBridge, R. B., Kirchner, J., Fearon, K., Mao, J. and Corcoran, K. (2000). "Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays." *Nat Biotechnol* **18**(6): 630-634.

Brzeski, J. and Jerzmanowski, A. (2003). "Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors." *J Biol Chem* **278**(2): 823-828.

Brzeski, J. and Jerzmanowski, A. (2004). "Plant chromatin -- epigenetics linked to ATP-dependent remodeling and architectural proteins." *FEBS Lett* **567**(1): 15-19.

Bühler, M., Verdel, A. and Moazed, D. (2006). "Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing." *Cell* **125**(5): 873-886.

Bushnell, D. A., Cramer, P. and Kornberg, R. D. (2002). "Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution." *Proc Natl Acad Sci U S A* **99**(3): 1218-1222.

Cam, H. P., Sugiyama, T., Chen, E. S., Chen, X., FitzGerald, P. C. and Grewal, S. I. (2005). "Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome." *Nat Genet* **37**(8): 809-819.

Campell, B. R., Song, Y., Posch, T. E., Cullis, C. A. and Town, C. D. (1992). "Sequence and organization of 5S ribosomal RNA-encoding genes of Arabidopsis thaliana." *Gene* **112**(2): 225-228.

Cao, X. and Jacobsen, S. E. (2002). "Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes." *Proc Natl Acad Sci U S A* **99 Suppl 4**: 16491-16498.

Cao, X. and Jacobsen, S. E. (2002). "Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing." *Curr Biol* **12**(13): 1138-1144.

Cao, X., Aufsatz, W., Zilberman, D., Mette, M. F., Huang, M. S., Matzke, M. and Jacobsen, S. E. (2003). "Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation." *Curr Biol* **13**(24): 2212-2217.

Cao, X., Springer, N. M., Muszynski, M. G., Phillips, R. L., Kaeppler, S. and Jacobsen, S. E. (2000). "Conserved plant genes with similarity to mammalian de novo DNA methyltransferases." *Proc Natl Acad Sci U S A* **97**(9): 4979-4984.

Carmell, M. A. and Hannon, G. J. (2004). "RNase III enzymes and the initiation of gene silencing." *Nat Struct Mol Biol* **11**(3): 214-218.

Carmell, M. A., Xuan, Z., Zhang, M. Q. and Hannon, G. J. (2002). "The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis." *Genes Dev* **16**(21): 2733-2742.

Carmo-Fonseca, M., Mendes-Soares, L. and Campos, I. (2000). "To be or not to be in the nucleolus." *Nat Cell Biol* **2**(6): E107-112.

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P. and Workman, J. L. (2005). "Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription." *Cell* **123**(4): 581-592.

Catalanotto, C., Pallotta, M., ReFalo, P., Sachs, M. S., Vayssie, L., Macino, G. and Cogoni, C. (2004). "Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in Neurospora crassa." *Mol Cell Biol* **24**(6): 2536-2545.

Caudy, A. A., Myers, M., Hannon, G. J. and Hammond, S. M. (2002). "Fragile X-related protein and VIG associate with the RNA interference machinery." *Genes Dev* **16**(19): 2491-2496.

Cervoni, N. and Szyf, M. (2001). "Demethylase activity is directed by histone acetylation." *J Biol Chem* **276**(44): 40778-40787.

Chan, S. W., Henderson, I. R., Zhang, X., Shah, G., Chien, J. S. and Jacobsen, S. E. (2006). "RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in arabidopsis." *PLoS Genet* **2**(6): e83.

Chan, S. W., Zilberman, D., Xie, Z., Johansen, L. K., Carrington, J. C. and Jacobsen, S. E. (2004). "RNA silencing genes control de novo DNA methylation." *Science* **303**(5662): 1336.

Chen, X., Liu, J., Cheng, Y. and Jia, D. (2002). "HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower." *Development* **129**(5): 1085-1094.

Chen, Z. J. and Pikaard, C. S. (1997). "Epigenetic silencing of RNA polymerase I transcription: a role for DNA methylation and histone modification in nucleolar dominance." *Genes Dev* **11**(16): 2124-2136.

Chen, Z. J. and Pikaard, C. S. (1997). "Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in Brassica." *Proc Natl Acad Sci U S A* **94**(7): 3442-3447.

Chen, Z. J., Comai, L. and Pikaard, C. S. (1998). "Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in Arabidopsis allopolyploids." *Proc Natl Acad Sci U S A* **95**(25): 14891-14896.

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K. and Shiekhattar, R. (2005). "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." *Nature* **436**(7051): 740-744.

Clough, S. J. and Bent, A. F. (1998). "Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana." *Plant J* **16**(6): 735-743.

Copenhaver, G. P. and Pikaard, C. S. (1996). "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**(2): 259-272.

Copenhaver, G. P. and Pikaard, C. S. (1996). "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**(2): 273-282.

Cremer, T. and Cremer, C. (2001). "Chromosome territories, nuclear architecture and gene regulation in mammalian cells." *Nat Rev Genet* **2**(4): 292-301.

Csordas, A. (1990). "On the biological role of histone acetylation." *Biochem J* **265**(1): 23-38.

Dahmus, M. E. (1996). "Reversible phosphorylation of the C-terminal domain of RNA polymerase II." *J Biol Chem* **271**(32): 19009-19012.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S. and Baulcombe, D. C. (2000). "An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus." *Cell* **101**(5): 543-553.

Dammann, R., Lucchini, R., Koller, T. and Sogo, J. M. (1993). "Chromatin structures and transcription of rDNA in yeast Saccharomyces cerevisiae." *Nucleic Acids Res* **21**(10): 2331-2338.

Dernburg, A. F., Sedat, J. W. and Hawley, R. S. (1996). "Direct evidence of a role for heterochromatin in meiotic chromosome segregation." *Cell* **86**(1): 135-146.

Djupedal, I., Portoso, M., Spahr, H., Bonilla, C., Gustafsson, C. M., Allshire, R. C. and Ekwall, K. (2005). "RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing." *Genes Dev* **19**(19): 2301-2306.

Doelling, J. H. and Pikaard, C. S. (1995). "The minimal ribosomal RNA gene promoter of Arabidopsis thaliana includes a critical element at the transcription initiation site." *Plant J* **8**(5): 683-692.

Doelling, J. H. and Pikaard, C. S. (1996). "Species-specificity of rRNA gene transcription in plants manifested as a switch in RNA polymerase specificity." *Nucleic Acids Res* **24**(23): 4725-4732.

Doelling, J. H., Gaudino, R. J. and Pikaard, C. S. (1993). "Functional analysis of Arabidopsis thaliana rRNA gene and spacer promoters in vivo and by transient expression." *Proc Natl Acad Sci U S A* **90**(16): 7528-7532.

Dover, G. (1982). "A role for the genome in the origin of species?" *Prog Clin Biol Res* **96**: 435-459.

Dover, G. A. and Flavell, R. B. (1984). "Molecular coevolution: DNA divergence and the maintenance of function." *Cell* **38**(3): 622-623.

Dutnall, R. N. (2003). "Cracking the histone code: one, two, three methyls, you're out!" *Mol Cell* **12**(1): 3-4.

Earley, K., Lawrence, R. J., Pontes, O., Reuther, R., Enciso, A. J., Silva, M., Neves, N., Gross, M., Viegas, W. and Pikaard, C. S. (2006). "Erasure of histone acetylation by Arabidopsis HDA6 mediates large-scale gene silencing in nucleolar dominance." *Genes Dev* **20**(10): 1283-1293.

Eberharter, A. and Becker, P. B. (2002). "Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics." *EMBO Rep* **3**(3): 224-229.

Eissenberg, J. C. and Elgin, S. C. (2000). "The HP1 protein family: getting a grip on chromatin." *Curr Opin Genet Dev* **10**(2): 204-210.

Ekwall, K., Javerzat, J. P., Lorentz, A., Schmidt, H., Cranston, G. and Allshire, R. (1995). "The chromodomain protein Swi6: a key component at fission yeast centromeres." *Science* **269**(5229): 1429-1431.

Fagard, M., Boutet, S., Morel, J. B., Bellini, C. and Vaucheret, H. (2000). "AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals." *Proc Natl Acad Sci U S A* **97**(21): 11650-11654.

Finnegan, E. J. and Kovac, K. A. (2000). "Plant DNA methyltransferases." *Plant Mol Biol* **43**(2-3): 189-201.

Finnegan, E. J., Genger, R. K., Peacock, W. J. and Dennis, E. S. (1998). "DNA Methylation in Plants." *Annu Rev Plant Physiol Plant Mol Biol* **49**: 223-247.

Finnegan, E. J., Peacock, W. J. and Dennis, E. S. (2000). "DNA methylation, a key regulator of plant development and other processes." *Curr Opin Genet Dev* **10**(2): 217-223.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S. E. and Mello, C. C. (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. " *Nature* **391**:806-811.

Flavell, R. B. (1986). "Repetitive DNA and chromosome evolution in plants." *Philos Trans R Soc Lond B Biol Sci* **312**(1154): 227-242.

Fodor, B. D., Kubicek, S., Yonezawa, M., O'Sullivan, R. J., Sengupta, R., Perez-Burgos, L., Opravil, S., Mechtler, K., Schotta, G. and Jenuwein, T. (2006). "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells." *Genes Dev* **20**(12): 1557-1562.

Fourel, G., Lebrun, E. and Gilson, E. (2002). "Protosilencers as building blocks for heterochromatin." *Bioessays* **24**(9): 828-835.

Francis, N. J. and Kingston, R. E. (2001). "Mechanisms of transcriptional memory." *Nat Rev Mol Cell Biol* **2**(6): 409-421.

Fransz, P., De Jong, J. H., Lysak, M., Castiglione, M. R. and Schubert, I. (2002). "Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate." *Proc Natl Acad Sci U S A* **99**(22): 14584-14589.

Fransz, P., Soppe, W. and Schubert, I. (2003). "Heterochromatin in interphase nuclei of Arabidopsis thaliana." *Chromosome Res* **11**(3): 227-240.

French, S. L., Osheim, Y. N., Cioci, F., Nomura, M. and Beyer, A. L. (2003). "In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes." *Mol Cell Biol* **23**(5): 1558-1568.

Gao, M. J., Schafer, U. A., Parkin, I. A., Hegedus, D. D., Lydiate, D. J. and Hannoufa, A. (2003). "A novel protein from Brassica napus has a putative KID domain and responds to low temperature." *Plant J* **33**(6): 1073-1086.

Garcia-Ramirez, M., Rocchini, C. and Ausio, J. (1995). "Modulation of chromatin folding by histone acetylation." *J Biol Chem* **270**(30): 17923-17928.

Gasciolli, V., Mallory, A. C., Bartel, D. P. and Vaucheret, H. (2005). "Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs." *Curr Biol* **15**(16): 1494-1500.

Gerbi, S. A. (1986). "The evolution of eukaryotic ribosomal DNA." *Biosystems* **19**(4): 247-258.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E. M. and Coupland, G. (1997). "A Polycomb-group gene regulates homeotic gene expression in Arabidopsis." *Nature* **386**(6620): 44-51.

Grewal, S. I. and Elgin, S. C. (2002). "Heterochromatin: new possibilities for the inheritance of structure." *Curr Opin Genet Dev* **12**(2): 178-187.

Grewal, S. I. and Jia, S. (2007). "Heterochromatin revisited." *Nat Rev Genet* **8**(1): 35-46.

Grewal, S. I. and Moazed, D. (2003). "Heterochromatin and epigenetic control of gene expression." *Science* **301**(5634): 798-802.

Grewal, S. I. and Rice, J. C. (2004). "Regulation of heterochromatin by histone methylation and small RNAs." *Curr Opin Cell Biol* **16**(3): 230-238.

Grishok, A. and Mello, C. C. (2002). "RNAi (Nematodes: Caenorhabditis elegans)." *Adv Genet* **46**: 339-360.

Grishok, A., Pasquinelli, A. E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G. and Mello, C. C. (2001). "Genes and mechanisms related to RNA interference regulate expression of

the small temporal RNAs that control C. elegans developmental timing." *Cell* **106**(1): 23-34.

Gruenbaum, Y., Naveh-Many, T., Cedar, H. and Razin, A. (1981). "Sequence specificity of methylation in higher plant DNA." *Nature* **292**(5826): 860-862.

Grummt, I. (2003). "Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus." *Genes Dev* **17**(14): 1691-1702.

Grummt, I. and Pikaard, C. S. (2003). "Epigenetic silencing of RNA polymerase I transcription." *Nat Rev Mol Cell Biol* **4**(8): 641-649.

Grummt, I., Roth, E. and Paule, M. R. (1982). "Ribosomal RNA transcription in vitro is species specific." *Nature* **296**(5853): 173-174.

Grunstein, M. (1997). "Histone acetylation in chromatin structure and transcription." *Nature* **389**(6649): 349-352.

Grunstein, M. (1998). "Yeast heterochromatin: regulation of its assembly and inheritance by histones." *Cell* **93**(3): 325-328.

Guseinov, V. A. and Vanyushin, B. F. (1975). "Content and localisation of 5-methylcytosine in DNA of healthy and wilt-infected cotton plants." *Biochim Biophys Acta* **395**(3): 229-238.

Gustafson, A. M., Allen, E., Givan, S., Smith, D., Carrington, J. C. and Kasschau, K. D. (2005). "ASRP: the Arabidopsis Small RNA Project Database." *Nucleic Acids Res* **33**(Database issue): D637-640.

Hall, I. M., Noma, K. and Grewal, S. I. (2003). "RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast." *Proc Natl Acad Sci U S A* **100**(1): 193-198.

Hall, I. M., Shankaranarayana, G. D., Noma, K., Ayoub, N., Cohen, A. and Grewal, S. I. (2002). "Establishment and maintenance of a heterochromatin domain." *Science* **297**(5590): 2232-2237.

Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R. and Hannon, G. J. (2001). "Argonaute2, a link between genetic and biochemical analyses of RNAi." *Science* **293**(5532): 1146-1150.

Hampsey, M. and Reinberg, D. (2003). "Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation." *Cell* **113**(4): 429-432.

Hanfstingl, U., Berry, A., Kellogg, E. A., Costa, J. T., 3rd, Rudiger, W. and Ausubel, F. M. (1994). "Haplotypic divergence coupled with lack of diversity at the Arabidopsis thaliana alcohol dehydrogenase locus: roles for both balancing and directional selection?" *Genetics* **138**(3): 811-828.

Hannan, K. M., Hannan, R. D. and Rothblum, L. I. (1998). "Transcription by RNA polymerase I." *Front Biosci* **3**: d376-398.

Hannon, G. J. (2002). "RNA interference." *Nature* **418**(6894): 244-251.

He, L., Thomson, J. M., Hemann, M. T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S. W., Hannon, G. J. and Hammond, S. M. (2005). "A microRNA polycistron as a potential human oncogene." *Nature* **435**(7043): 828-833.

He, Z. and Sontheimer, E. J. (2004). ""siRNAs and miRNAs": a meeting report on RNA silencing." *Rna* **10**(8): 1165-1173.

Heard, E. (2005). "Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome." *Curr Opin Genet Dev* **15**(5): 482-489.

Heitz, E. (1928). "Das heterochromatin der Moose". *Jahrb Wiss Botanik* **69**:762-818. (*cit in* Pikaard, 2000).

Heitz, E. (1931). "Die Ursache der gesetzmaszigen Zahl, Lage, Form und Grosze pflanzlicher Nukleolen." *Planta* **12**: 775-844. (*cit in* Pikaard, 2000)

Helliwell, C. A., Peacock, W. J. and Dennis, E. S. (2002). "Isolation and functional characterization of cytochrome P450s in gibberellin biosynthesis pathway." *Methods Enzymol* **357**: 381-388.

Henikoff, S. (2000). "Heterochromatin function in complex genomes." *Biochim Biophys Acta* **1470**(1): O1-8.

Henikoff, S. and Comai, L. (1998). "A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis." *Genetics* **149**(1): 307-318.

Herr, A. J., Jensen, M. B., Dalmay, T. and Baulcombe, D. C. (2005). "RNA polymerase IV directs silencing of endogenous DNA." *Science* **308**(5718): 118-120.

Heslop-Harrison, J. S. (1998). "Cytogenetic analysis of Arabidopsis." *Methods Mol Biol* **82**: 119-127.

Heslop-Harrison, J. S., Brandes, A. and Schwarzacher, T. (2003). "Tandemly repeated DNA sequences and centromeric chromosomal regions of Arabidopsis species." *Chromosome Res* **11**(3): 241-253.

Horn, P.J., Peterson, C.L. (2002). "Chromatin higher order folding: Wrapping up transcription." *Science* **297**:1824-1827.

Hsu, C. L. and Stevens, A. (1993). "Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure." *Mol Cell Biol* **13**(8): 4826-4835.

Jackson, J. P., Johnson, L., Jasencakova, Z., Zhang, X., PerezBurgos, L., Singh, P. B., Cheng, X., Schubert, I., Jenuwein, T. and Jacobsen, S. E. (2004). "Dimethylation of histone H3 lysine 9 is a critical mark for DNA methylation and gene silencing in Arabidopsis thaliana." *Chromosoma* **112**(6): 308-315.

Jackson, J. P., Lindroth, A. M., Cao, X. and Jacobsen, S. E. (2002). "Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase." *Nature* **416**(6880): 556-560.

Jeddeloh, J. A., Stokes, T. L. and Richards, E. J. (1999). "Maintenance of genomic methylation requires a SWI2/SNF2-like protein." *Nat Genet* **22**(1): 94-97.

Jenuwein, T. and Allis, C. D. (2001). "Translating the histone code." *Science* **293**(5532): 1074-1080.

Jeppesen, P. and Turner, B. M. (1993). "The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression." *Cell* **74**(2): 281-289.

Joshi, A. A. and Struhl, K. (2005). "Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation." *Mol Cell* **20**(6): 971-978.

Kakutani, T., Jeddeloh, J. A. and Richards, E. J. (1995). "Characterization of an Arabidopsis thaliana DNA hypomethylation mutant." *Nucleic Acids Res* **23**(1): 130-137.

Kamm, A., Galasso, I., Schmidt, T. and Heslop-Harrison, J. S. (1995). "Analysis of a repetitive DNA family from Arabidopsis arenosa and relationships between Arabidopsis species." *Plant Mol Biol* **27**(5): 853-862.

Kanno, T., Huettel, B., Mette, M. F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D. P., Matzke, M. and Matzke, A. J. (2005). "Atypical RNA polymerase subunits required for RNA-directed DNA methylation." *Nat Genet* **37**(7): 761-765.

Kanno, T., Huettel, B., Mette, M. F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D. P., Matzke, M. and Matzke, A. J. (2005). "Atypical RNA polymerase subunits required for RNA-directed DNA methylation." *Nat Genet* **37**(7): 761-765.

Kanno, T., Mette, M. F., Kreil, D. P., Aufsatz, W., Matzke, M. and Matzke, A. J. (2004). "Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation." *Curr Biol* **14**(9): 801-805.

Karpen, G. H. and Allshire, R. C. (1997). "The case for epigenetic effects on centromere identity and function." *Trends Genet* **13**(12): 489-496.

Kastenmayer, J. P. and Green, P. J. (2000). "Novel features of the XRN-family in Arabidopsis: evidence that AtXRN4, one of several orthologs of nuclear Xrn2p/Rat1p, functions in the cytoplasm." *Proc Natl Acad Sci U S A* **97**(25): 13985-13990.

Kato, H., Goto, D. B., Martienssen, R. A., Urano, T., Furukawa, K. and Murakami, Y. (2005). "RNA polymerase II is required for RNAi-dependent heterochromatin assembly." *Science* **309**(5733): 467-469.

Kellum, R. and Alberts, B. M. (1995). "Heterochromatin protein 1 is required for correct chromosome segregation in Drosophila embryos." *J Cell Sci* **108 (Pt 4)**: 1419-1431.

Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. and Plasterk, R. H. (2001). "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans." *Genes Dev* **15**(20): 2654-2659.

Khusial, P. R., Vaidya, K. and Zieve, G. W. (2005). "The symmetrical dimethylarginine post-translational modification of the SmD3 protein is not required for snRNP assembly and nuclear transport." *Biochem Biophys Res Commun* **337**(4): 1119-1124.

Kidner, C. A. and Martienssen, R. A. (2004). "Spatially restricted microRNA directs leaf polarity through ARGONAUTE1." *Nature* **428**(6978): 81-84.

Kidner, C. A. and Martienssen, R. A. (2005). "The developmental role of microRNA in plants." *Curr Opin Plant Biol* **8**(1): 38-44.

Knight, S. W. and Bass, B. L. (2001). "A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans." *Science* **293**(5538): 2269-2271.

Kurihara, Y. and Watanabe, Y. (2004). "Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions." *Proc Natl Acad Sci U S A* **101**(34): 12753-12758.

Lachner, M. and Jenuwein, T. (2002). "The many faces of histone lysine methylation." *Curr Opin Cell Biol* **14**(3): 286-298.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001). "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins." *Nature* **410**(6824): 116-120.

Lawrence, R. J. and Pikaard, C. S. (2003). "Transgene-induced RNA interference: a strategy for overcoming gene redundancy in polyploids to generate loss-of-function mutations." *Plant J* **36**(1): 114-121.

Lawrence, R. J., Earley, K., Pontes, O., Silva, M., Chen, Z. J., Neves, N., Viegas, W. and Pikaard, C. S. (2004). "A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance." *Mol Cell* **13**(4): 599-609.

Le Hir, H., Izaurralde, E., Maquat, L. E. and Moore, M. J. (2000). "The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions." *Embo J* **19**(24): 6860-6869.

Lee, H. S. and Chen, Z. J. (2001). "Protein-coding genes are epigenetically regulated in Arabidopsis polyploids." *Proc Natl Acad Sci U S A* **98**(12): 6753-6758.

Lee, M. G., Wynder, C., Cooch, N. and Shiekhattar, R. (2005). "An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation." *Nature* **437**(7057): 432-435.

Lee, Y. S., Nakahara, K., Pham, J. W., Kim, K., He, Z., Sontheimer, E. J. and Carthew, R. W. (2004). "Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways." *Cell* **117**(1): 69-81.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). "Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin." *Curr Biol* **13**(14): 1192-1200.

Lewis, M. S. and Pikaard, C. S. (2001). "Restricted chromosomal silencing in nucleolar dominance." *Proc Natl Acad Sci U S A* **98**(25): 14536-14540.

Lewis, M. S., Cheverud, J. M. and Pikaard, C. S. (2004). "Evidence for nucleolus organizer regions as the units of regulation in nucleolar dominance in Arabidopsis thaliana interecotype hybrids." *Genetics* **167**(2): 931-939.

Li, C. F., Pontes, O., El-Shami, M., Henderson, I. R., Bernatavichute, Y. V., Chan, S. W., Lagrange, T., Pikaard, C. S. and Jacobsen, S. E. (2006). "An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana." *Cell* **126**(1): 93-106.

Li, J., Langst, G. and Grummt, I. (2006). "NoRC-dependent nucleosome positioning silences rRNA genes." *Embo J* **25**(24): 5735-5741.

Li, J., Yang, Z., Yu, B., Liu, J. and Chen, X. (2005). "Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis." *Curr Biol* **15**(16): 1501-1507.

Lin, X., Kaul, S., Rounsley, S., Shea, T. P., Benito, M. I., Town, C. D., Fujii, C. Y., Mason, T., Bowman, C. L., Barnstead, M., Feldblyum, T. V., Buell, C. R., Ketchum, K. A., Lee, J., Ronning, C. M., Koo, H. L., Moffat, K. S., Cronin, L. A., Shen, M., Pai, G., Van Aken, S., Umayam, L., Tallon, L. J., Gill, J. E., Adams, M. D., Carrera, A. J., Creasy, T. H., Goodman, H. M., Somerville, C. R., Copenhaver, G. P., Preuss, D., Nierman, W. C., White, O., Eisen, J. A., Salzberg, S. L., Fraser, C. M. and Venter, J. C. (1999). "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana." *Nature* **402**(6763): 761-768.

Lindroth, A. M., Cao, X., Jackson, J. P., Zilberman, D., McCallum, C. M., Henikoff, S. and Jacobsen, S. E. (2001). "Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation." *Science* **292**(5524): 2077-2080.

Lippman, Z., Gendrel, A. V., Black, M., Vaughn, M. W., Dedhia, N., McCombie, W. R., Lavine, K., Mittal, V., May, B., Kasschau, K. D., Carrington, J. C., Doerge, R. W., Colot, V. and Martienssen, R. (2004). "Role of transposable elements in heterochromatin and epigenetic control." *Nature* **430**(6998): 471-476.

Lippman, Z., May, B., Yordan, C., Singer, T. and Martienssen, R. (2003). "Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification." *PLoS Biol* **1**(3): E67.

Liu, J., Carmell, M. A., Rivas, F. V., Marsden, C. G., Thomson, J. M., Song, J. J., Hammond, S. M., Joshua-Tor, L. and Hannon, G. J. (2004). "Argonaute2 is the catalytic engine of mammalian RNAi." *Science* **305**(5689): 1437-1441.

Liu, J., Rivas, F. V., Wohlschlegel, J., Yates, J. R., 3rd, Parker, R. and Hannon, G. J. (2005). "A role for the P-body component GW182 in microRNA function." *Nat Cell Biol* **7**(12): 1261-1266.

Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. and Parker, R. (2005). "MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies." *Nat Cell Biol* **7**(7): 719-723.

Llave, C., Kasschau, K. D., Rector, M. A. and Carrington, J. C. (2002). "Endogenous and silencing-associated small RNAs in plants." *Plant Cell* **14**(7): 1605-1619.

Loidl, P. (1994). "Histone acetylation: facts and questions." *Chromosoma* **103**(7): 441-449.

Loidl, P. (2004). "A plant dialect of the histone language." *Trends Plant Sci* **9**(2): 84-90.

Lu, C., Tej, S. S., Luo, S., Haudenschild, C. D., Meyers, B. C. and Green, P. J. (2005). "Elucidation of the small RNA component of the transcriptome." *Science* **309**(5740): 1567-1569.

Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. and Richmond, T. J. (1997). "Characterization of nucleosome core particles containing histone proteins made in bacteria." *J Mol Biol* **272**(3): 301-311.

Lykke-Andersen, J., Shu, M. D. and Steitz, J. A. (2000). "Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon." *Cell* **103**(7): 1121-1131.

Maine, E. M. (2001). "RNAi as a tool for understanding germline development in Caenorhabditis elegans: Uses and cautions." *Developmental Biology* **239**: 177-189.

Makeyev, E. V. and Bamford, D. H. (2002). "Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes." *Mol Cell* **10**(6): 1417-1427.

Malagnac, F., Bartee, L. and Bender, J. (2002). "An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation." *Embo J* **21**(24): 6842-6852.

Maquat, L. E. (2005). "Nonsense-mediated mRNA decay in mammals." *J Cell Sci* **118**(Pt 9): 1773-1776.

Martinez-Zapater, J. M., Estelle, M.A., Somerville, C.R. (1986). "A highly repeated DNA sequence in Arabidopsis thaliana." *Mol Gen Genet* **204**: 417-423.

Mathieu, O., Jasencakova, Z., Vaillant, I., Gendrel, A. V., Colot, V., Schubert, I. and Tourmente, S. (2003). "Changes in 5S rDNA chromatin organization and transcription during heterochromatin establishment in Arabidopsis." *Plant Cell* **15**(12): 2929-2939.

Matzke, M. A. and Birchler, J. A. (2005). "RNAi-mediated pathways in the nucleus." *Nat Rev Genet* **6**(1): 24-35.

May, B. P., Lippman, Z. B., Fang, Y., Spector, D. L. and Martienssen, R. A. (2005). "Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats." *PLoS Genet* **1**(6): e79.

Mayer, C., Schmitz, K. M., Li, J., Grummt, I. and Santoro, R. (2006). "Intergenic transcripts regulate the epigenetic state of rRNA genes." *Mol Cell* **22**(3): 351-361.

Mayer, K., Schuller, C., Wambutt, R., Murphy, G., Volckaert, G., Pohl, T., Dusterhoft, A., Stiekema, W., Entian, K. D., Terryn, N., Harris, B., Ansorge, W., Brandt, P., Grivell, L., Rieger, M., Weichselgartner, M., de Simone, V., Obermaier, B., Mache, R., Muller, M., Kreis, M., Delseny, M., Puigdomenech, P., Watson, M., Schmidtheini, T., Reichert, B., Portatelle, D., Perez-Alonso, M., Boutry, M., Bancroft, I., Vos, P., Hoheisel, J., Zimmermann, W., Wedler, H., Ridley, P., Langham, S. A., McCullagh, B., Bilham, L., Robben, J., Van der Schueren, J., Grymonprez, B., Chuang, Y. J., Vandenbussche, F., Braeken, M., Weltjens, I., Voet, M., Bastiaens, I., Aert, R., Defoor, E., Weitzenegger, T., Bothe, G., Ramsperger, U., Hilbert, H., Braun, M., Holzer, E., Brandt, A., Peters, S., van Staveren, M., Dirske, W., Mooijman, P., Klein Lankhorst, R., Rose, M., Hauf, J., Kotter, P., Berneiser, S., Hempel, S., Feldpausch, M., Lamberth, S., Van den Daele, H., De Keyser, A., Buysshaert, C., Gielen, J., Villarroel, R., De Clercq, R., Van Montagu, M., Rogers, J., Cronin, A., Quail, M., Bray-Allen, S., Clark, L., Doggett, J., Hall, S., Kay, M., Lennard, N., McLay, K., Mayes, R., Pettett, A., Rajandream, M. A., Lyne, M., Benes, V., Rechmann, S., Borkova, D., Blocker, H., Scharfe, M., Grimm, M., Lohnert, T. H., Dose, S., de Haan, M., Maarse, A., Schafer, M., Muller-Auer, S., Gabel, C., Fuchs, M., Fartmann, B., Granderath, K., Dauner, D., Herzl, A., Neumann, S., Argiriou, A., Vitale, D., Liguori, R., Piravandi, E., Massenet, O., Quigley, F., Clabauld, G., Mundlein, A., Felber, R., Schnabl, S., Hiller, R., Schmidt, W., Lecharny, A., Aubourg, S., Chefdor, F., Cooke, R., Berger, C., Montfort, A., Casacuberta, E., Gibbons, T., Weber, N., Vandenbol, M., Bargues, M., Terol, J., Torres, A., Perez-Perez, A., Purnelle, B., Bent, E., Johnson, S., Tacon, D., Jesse, T., Heijnen, L., Schwarz, S., Scholler, P., Heber, S., Francs, P., Bielke, C., Frishman, D., Haase, D., Lemcke, K., Mewes, H. W., Stocker, S., Zaccaria, P., Bevan, M., Wilson, R. K., de la Bastide, M., Habermann, K., Parnell, L., Dedhia, N., Gnoj, L., Schutz, K., Huang, E., Spiegel, L., Sehkon, M., Murray, J., Sheet, P., Cordes, M., Abu-Threideh, J., Stoneking, T., Kalicki, J., Graves, T., Harmon, G., Edwards, J., Latreille, P., Courtney, L., Cloud, J., Abbott, A., Scott, K., Johnson, D., Minx, P., Bentley, D., Fulton, B., Miller, N., Greco, T., Kemp, K., Kramer, J., Fulton, L., Mardis, E., Dante, M., Pepin, K., Hillier, L., Nelson, J., Spieth, J., Ryan, E., Andrews, S., Geisel, C., Layman, D., Du, H., Ali, J., Berghoff, A., Jones, K., Drone, K., Cotton, M., Joshu, C., Antonoiu, B., Zidanic, M., Strong, C., Sun, H., Lamar, B., Yordan, C., Ma, P., Zhong, J., Preston, R., Vil, D., Shekher, M.,

Matero, A., Shah, R., Swaby, I. K., O'Shaughnessy, A., Rodriguez, M., Hoffmann, J., Till, S., Granat, S., Shohdy, N., Hasegawa, A., Hameed, A., Lodhi, M., Johnson, A., Chen, E., Marra, M., Martienssen, R. and McCombie, W. R. (1999). "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana." *Nature* **402**(6763): 769-777.

McClintock, B. (1934). "The relation of a particular chromosomal element to the development of the nucleoli in Zea mays Zeit." *Zellforsch mik Anat* **21**: 294-328.

Melese, T. and Xue, Z. (1995). "The nucleolus: an organelle formed by the act of building a ribosome." *Curr Opin Cell Biol* **7**(3): 319-324.

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A. and Matzke, A. J. (2000). "Transcriptional silencing and promoter methylation triggered by double-stranded RNA." *Embo J* **19**(19): 5194-5201.

Mette, M. F., Kanno, T., Aufsatz, W., Jakowitsch, J., van der Winden, J., Matzke, M. A. and Matzke, A. J. (2002). "Endogenous viral sequences and their potential contribution to heritable virus resistance in plants." *Embo J* **21**(3): 461-469.

Metzger, E., Wissmann, M., Yin, N., Muller, J. M., Schneider, R., Peters, A. H., Gunther, T., Buettner, R. and Schule, R. (2005). "LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription." *Nature* **437**(7057): 436-439.

Meyer, P., Niedenhof, I. and ten Lohuis, M. (1994). "Evidence for cytosine methylation of non-symmetrical sequences in transgenic Petunia hybrida." *Embo J* **13**(9): 2084-2088.

Meyerowitz, E. M. (2002). "Plants compared to animals: the broadest comparative study of development." *Science* **295**(5559): 1482-1485.

Mito, Y., Henikoff, J. G. and Henikoff, S. (2005). "Genome-scale profiling of histone H3.3 replacement patterns." *Nat Genet* **37**(10): 1090-1097.

Morales-Ruiz, T., Ortega-Galisteo, A. P., Ponferrada-Marin, M. I., Martinez-Macias, M. I., Ariza, R. R. and Roldan-Arjona, T. (2006). "DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases." *Proc Natl Acad Sci U S A* **103**(18): 6853-6858.

Morel, J. B., Godon, C., Mourrain, P., Beclin, C., Boutet, S., Feuerbach, F., Proux, F. and Vaucheret, H. (2002). "Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance." *Plant Cell* **14**(3): 629-639.

Moss, T. and Birnstiel, M. L. (1979). "The putative promoter of a Xenopus laevis ribosomal gene is reduplicated." *Nucleic Acids Res* **6**(12): 3733-3743.

Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P. and Moazed, D. (2004). "Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs." *Cell* **119**(6): 789-802.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Jouette, D., Lacombe, A. M., Nikic, S., Picault, N., Remoue, K., Sanial, M., Vo, T. A. and Vaucheret, H. (2000). "Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance." *Cell* **101**(5): 533-542.

Muhlrad, D. and Parker, R. (1994). "Premature translational termination triggers mRNA decapping." *Nature* **370**(6490): 578-581.

Murfett, J., Wang, X. J., Hagen, G. and Guilfoyle, T. J. (2001). "Identification of Arabidopsis histone deacetylase HDA6 mutants that affect transgene expression." *Plant Cell* **13**(5): 1047-1061.

Murray, K. (1964). "The Occurrence of Epsilon-N-Methyl Lysine in Histones." *Biochemistry* **3**: 10-15. (*cit in* Bannister *et al.*, 2002).

Muscarella, D. E., Vogt, V. M. and Bloom, S. E. (1985). "The ribosomal RNA gene cluster in aneuploid chickens: evidence for increased gene dosage and regulation of gene expression." *J Cell Biol* **101**(5 Pt 1): 1749-1756.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D. and Grewal, S. I. (2001). "Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly." *Science* **292**(5514): 110-113.

Navashin, M. (1934). "Chromosome alterations caused by hybridization and their bearing upon certain general genetic problems." *Cytologia* **5**: 169-203. (*cit in* Pikaard, 2000)

Neves, N., Silva, M., Heslop-Harrison, J. S. and Viegas, W. (1997). "Nucleolar dominance in triticales: control by unlinked genes." *Chromosome Res* **5**(2): 125-131.

Nonaka, K., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S. and Watanabe, Y. (2001). "Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast". *Nat Cell Biol* **4**:89-93.

O'Kane, D., Gill, V., Boyd, P. and Burdon, R. (1996). "Chilling, oxidative stress and antioxidant responses in Arabidopsis thaliana callus." *Planta* **198**(3): 371-377.

Olson, M. O., Dundr, M. and Szebeni, A. (2000). "The nucleolus: an old factory with unexpected capabilities." *Trends Cell Biol* **10**(5): 189-196.

Onodera, Y., Haag, J. R., Ream, T., Nunes, P. C., Pontes, O. and Pikaard, C. S. (2005). "Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation." *Cell* **120**(5): 613-622.

Pandey, R., Muller, A., Napoli, C. A., Selinger, D. A., Pikaard, C. S., Richards, E. J., Bender, J., Mount, D. W. and Jorgensen, R. A. (2002). "Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes." *Nucleic Acids Res* **30**(23): 5036-5055.

Papp, I., Mette, M. F., Aufsatz, W., Daxinger, L., Schauer, S. E., Ray, A., van der Winden, J., Matzke, M. and Matzke, A. J. (2003). "Evidence for nuclear processing of plant micro RNA and short interfering RNA precursors." *Plant Physiol* **132**(3): 1382-1390.

Paule, M. R. and White, R. J. (2000). "SURVEY AND SUMMARY Transcription by RNA polymerases I and III." *Nucleic Acids Res* **28**: 1283-1298.

Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., Mann, M., Lamond, A. I., Brown, J. W. and Shaw, P. J. (2005). "Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions." *Mol Biol Cell* **16**(1): 260-269.

Penny, G. D., Kay, G. F., Sheardown, S. A., Rastan, S. and Brockdorff, N. (1996). "Requirement for Xist in X chromosome inactivation." *Nature* **379**(6561): 131-137.

Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H. L. and Poethig, R. S. (2004). "SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis." *Genes Dev* **18**(19): 2368-2379.

Peters, A. H., Kubicek, S., Mechtler, K., O'Sullivan, R. J., Derijck, A. A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., Martens, J. H. and Jenuwein, T. (2003). "Partitioning and plasticity of repressive histone methylation states in mammalian chromatin." *Mol Cell* **12**(6): 1577-1589.

Petes, T. D. (1980). "Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes." *Cell* **19**(3): 765-774.

Peyroche, G., Milkereit, P., Bischler, N., Tschochner, H., Schultz, P., Sentenac, A., Carles, C. and Riva, M. (2000). "The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3." *Embo J* **19**(20): 5473-5482.

Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. and Sontheimer, E. J. (2004). "A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila." *Cell* **117**(1): 83-94.

Phillips, D. M., Phillips, S.G. (1971). "Distinctive characteristics of nucleoli of two established cell lines." *J Cell Biol* **49**(3): 803-815.

Pikaard, C. S. (2000). "The epigenetics of nucleolar dominance." *Trends Genet* **16**(11): 495-500.

Pikaard, C. S. and Lawrence, R. J. (2002). "Uniting the paths to gene silencing." *Nat Genet* **32**(3): 340-341.

Pipal, A., Goralik-Schramel, M., Lusser, A., Lanzanova, C., Sarg, B., Loidl, A., Lindner, H., Rossi, V. and Loidl, P. (2003). "Regulation and processing of maize histone deacetylase Hda1 by limited proteolysis." *Plant Cell* **15**(8): 1904-1917.

Pontes, O., Lawrence, R. J., Neves, N., Silva, M., Lee, J. H., Chen, Z. J., Viegas, W. and Pikaard, C. S. (2003). "Natural variation in nucleolar dominance reveals the relationship between nucleolus organizer chromatin topology and rRNA gene transcription in Arabidopsis." *Proc Natl Acad Sci U S A* **100**(20): 11418-11423.

Pontes, O., Li, C. F., Nunes, P. C., Haag, J., Ream, T., Vitins, A., Jacobsen, S. E. and Pikaard, C. S. (2006). "The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center." *Cell* **126**(1): 79-92.

Pontes, O., Neves, N., Silva, M., Lewis, M. S., Madlung, A., Comai, L., Viegas, W. and Pikaard, C. S. (2004). "Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid Arabidopsis suecica genome." *Proc Natl Acad Sci U S A* **101**(52): 18240-18245.

Pontier, D., Yahubyan, G., Vega, D., Bulski, A., Saez-Vasquez, J., Hakimi, M. A., Lerbs-Mache, S., Colot, V. and Lagrange, T. (2005). "Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis." *Genes Dev* **19**(17): 2030-2040.

Price, R. A., Palmer, J.D., Al-Shehbaz, I.A. (1994). Systematic relationships of Arabidopsis: a molecular and morphological perspective. *Arabidopsis*. E. M. Meyerowitz, Somerville C.R., Cold Spring Harbor Laboratory Press, NY**:** 7–19.

Prior, C. P., Cantor, C. R., Johnson, E. M., Littau, V. C. and Allfrey, V. G. (1983). "Reversible changes in nucleosome structure and histone H3 accessibility in transcriptionally active and inactive states of rDNA chromatin." *Cell* **34**(3): 1033-1042.

Probst, A. V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., Lawrence, R. J., Pikaard, C. S., Murfett, J., Furner, I., Vaucheret, H. and Mittelsten Scheid, O. (2004). "Arabidopsis histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats." *Plant Cell* **16**(4): 1021-1034.

Qi, Y., Denli, A. M. and Hannon, G. J. (2005). "Biochemical specialization within Arabidopsis RNA silencing pathways." *Mol Cell* **19**(3): 421-428.

Qi, Y., He, X., Wang, X. J., Kohany, O., Jurka, J. and Hannon, G. J. (2006). "Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation." *Nature* **443**(7114): 1008-1012.

Reeder, R. H. (1974). *Ribosomes. Eds.* M. Nomura. NY, *Cold Spring Harbor.*489-519.

Reeder, R. H. (1989). "Regulatory elements of the generic ribosomal gene." *Curr Opin Cell Biol* **1**(3): 466-474.

Reinke, H. and Horz, W. (2003). "Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter." *Mol Cell* **11**(6): 1599-1607.

Rice, J. C. and Allis, C. D. (2001). "Histone methylation versus histone acetylation: new insights into epigenetic regulation." *Curr Opin Cell Biol* **13**(3): 263-273.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y. and Allis, C.D. (2003). "Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains." *Mol Cell* **12**(6):1591-1598.

Richards, E. J. and Elgin, S. C. (2002). "Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects." *Cell* **108**(4): 489-500.

Ridgway, P. and Almouzni, G. (2001). "Chromatin assembly and organization." *J Cell Sci* **114**(Pt 15): 2711-2712.

Rivin, C. J., Cullis, C. A. and Walbot, V. (1986). "Evaluating quantitative variation in the genome of Zea mays." *Genetics* **113**(4): 1009-1019.

Rossi, V., Locatelli, S., Lanzanova, C., Boniotti, M. B., Varotto, S., Pipal, A., Goralik-Schramel, M., Lusser, A., Gatz, C., Gutierrez, C. and Motto, M. (2003). "A maize histone deacetylase and retinoblastoma-related protein physically interact and cooperate in repressing gene transcription." *Plant Mol Biol* **51**(3): 401-413.

Round, E. K., Flowers, S. K. and Richards, E. J. (1997). "Arabidopsis thaliana centromere regions: genetic map positions and repetitive DNA structure." *Genome Res* **7**(11): 1045-1053.

Rountree, M. R. and Selker, E. U. (1997). "DNA methylation inhibits elongation but not initiation of transcription in Neurospora crassa." *Genes Dev* **11**(18): 2383-2395.

Russell, J. and Zomerdijk, J. C. (2005). "RNA-polymerase-I-directed rDNA transcription, life and works." *Trends Biochem Sci* **30**(2): 87-96.

Saez-Vasquez, J. and Pikaard, C. S. (1997). "Extensive purification of a putative RNA polymerase I holoenzyme from plants that accurately initiates rRNA gene transcription in vitro." *Proc Natl Acad Sci U S A* **94**(22): 11869-11874.

Sakashita, E., Tatsumi, S., Werner, D., Endo, H. and Mayeda, A. (2004). "Human RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator in vivo." *Mol Cell Biol* **24**(3): 1174-1187.

Sambrook, J., Russel D.W. (2001). *Molecular Cloning: A Laboratory Manual.*, *Eds.* Cold Spring Harbor, New York., *Cold Spring Harbor Laboratory Press*

Sandmeier, J. J., French, S., Osheim, Y., Cheung, W. L., Gallo, C. M., Beyer, A. L. and Smith, J. S. (2002). "RPD3 is required for the inactivation of yeast ribosomal DNA genes in stationary phase." *Embo J* **21**(18): 4959-4968.

Santoro, R. and Grummt, I. (2001). "Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription." *Mol Cell* **8**(3): 719-725.

Santoro, R. and Grummt, I. (2005). "Epigenetic mechanism of rRNA gene silencing: temporal order of NoRC-mediated histone modification, chromatin remodeling, and DNA methylation." *Mol Cell Biol* **25**(7): 2539-2546.

Santoro, R., Li, J. and Grummt, I. (2002). "The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription." *Nat Genet* **32**(3): 393-396.

Saze, H., Mittelsten Scheid, O. and Paszkowski, J. (2003). "Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis." *Nat Genet* **34**(1): 65-69.

Schauer, S. E., Jacobsen, S. E., Meinke, D. W. and Ray, A. (2002). "DICER-LIKE1: blind men and elephants in Arabidopsis development." *Trends Plant Sci* **7**(11): 487-491.

Scheer, U. and Hock, R. (1999). "Structure and function of the nucleolus." *Curr Opin Cell Biol* **11**(3): 385-390.

Scheer, U. and Weisenberger, D. (1994). "The nucleolus." *Curr Opin Cell Biol* **6**(3): 354-359.

Schiebel, W., Haas, B., Marinkovic, S., Klanner, A. and Sanger, H. L. (1993). "RNA-directed RNA polymerase from tomato leaves. II. Catalytic in vitro properties." *J Biol Chem* **268**(16): 11858-11867.

Schramke, V., Sheedy, D. M., Denli, A. M., Bonila, C., Ekwall, K., Hannon, G. J. and Allshire, R. C. (2005). "RNA-interference-directed chromatin modification coupled to RNA polymerase II transcription." *Nature* **435**(7046): 1275-1279.

Schwab, R., Ossowski, S., Riester, M., Warthmann, N. and Weigel, D. (2006). "Highly specific gene silencing by artificial microRNAs in Arabidopsis." *Plant Cell* **18**(5): 1121-1133.

Schwarz, D. S., Hutvagner, G., Haley, B. and Zamore, P. D. (2002). "Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways." *Mol Cell* **10**(3): 537-548.

Selker, E. U. (1998). "Trichostatin A causes selective loss of DNA methylation in Neurospora." *Proc Natl Acad Sci U S A* **95**(16): 9430-9435.

Shaw, P. J. and Jordan, E. G. (1995). "The nucleolus." *Annu Rev Cell Dev Biol* **11**: 93-121.

Shi, Y. J., Matson, C., Lan, F., Iwase, S., Baba, T. and Shi, Y. (2005). "Regulation of LSD1 histone demethylase activity by its associated factors." *Mol Cell* **19**(6): 857-864.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A. and Casero, R. A. (2004). "Histone demethylation mediated by the nuclear amine oxidase homolog LSD1." *Cell* **119**(7): 941-953.

Sigova, A., Rhind, N. and Zamore, P. D. (2004). "A single Argonaute protein mediates both transcriptional and posttranscriptional silencing in Schizosaccharomyces pombe." *Genes Dev* **18**(19): 2359-2367.

Silva, J. M., Hammond, S. M. and Hannon, G. J. (2002). "RNA interference: a promising approach to antiviral therapy?" *Trends Mol Med* **8**(11): 505-508.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J. and Plasterk, R. H. (2002). "Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi." *Curr Biol* **12**(15): 1317-1319.

Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N. and Maine, E. M. (2000). "EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans." *Curr Biol* **10**(4): 169-178.

Song, J. J., Smith, S. K., Hannon, G. J. and Joshua-Tor, L. (2004). "Crystal structure of Argonaute and its implications for RISC slicer activity." *Science* **305**(5689): 1434-1437.

Soppe, W. J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M. S., Jacobsen, S. E., Schubert, I. and Fransz, P. F. (2002). "DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis." *Embo J* **21**(23): 6549-6559.

Souret, F. F., Kastenmayer, J. P. and Green, P. J. (2004). "AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets." *Mol Cell* **15**(2): 173-183.

Spencer, V. A. and Davie, J. R. (1999). "Role of covalent modifications of histones in regulating gene expression." *Gene* **240**(1): 1-12.

Stevens, A., Hsu, C. L., Isham, K. R. and Larimer, F. W. (1991). "Fragments of the internal transcribed spacer 1 of pre-rRNA accumulate in Saccharomyces cerevisiae lacking 5'----3' exoribonuclease 1." *J Bacteriol* **173**(21): 7024-7028.

Strahl, B. D. and Allis, C. D. (2000). "The language of covalent histone modifications." *Nature* **403**(6765): 41-45.

Strohner, R., Nemeth, A., Jansa, P., Hofmann-Rohrer, U., Santoro, R., Langst, G. and Grummt, I. (2001). "NoRC--a novel member of mammalian ISWI-containing chromatin remodeling machines." *Embo J* **20**(17): 4892-4900.

Struhl, K. (1998). "Histone acetylation and transcriptional regulatory mechanisms." *Genes Dev* **12**(5): 599-606.

Szostak, J. W. and Wu, R. (1980). "Unequal crossing over in the ribosomal DNA of Saccharomyces cerevisiae." *Nature* **284**(5755): 426-430.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A. and Mello, C. C. (1999). "The rde-1 gene, RNA interference, and transposon silencing in C. elegans." *Cell* **99**(2): 123-132.

Tabara, H., Yigit, E., Siomi, H. and Mello, C. C. (2002). "The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans." *Cell* **109**(7): 861-871.

Tamaru, H. and Selker, E. U. (2003). "Synthesis of signals for de novo DNA methylation in Neurospora crassa." *Mol Cell Biol* **23**(7): 2379-2394.

Tijsterman, M. and Plasterk, R. H. (2004). "Dicers at RISC; the mechanism of RNAi." *Cell* **117**(1): 1-3.

Tran, R. K., Henikoff, J. G., Zilberman, D., Ditt, R. F., Jacobsen, S. E. and Henikoff, S. (2005). "DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes." *Curr Biol* **15**(2): 154-159.

Turner, B. M. (2000). "Histone acetylation and an epigenetic code." *Bioessays* **22**(9): 836-845.

Turner, B. M. (2002). "Cellular memory and the histone code." *Cell* **111**(3): 285-291.

Turner, B. M., Birley, A. J. and Lavender, J. (1992). "Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei." *Cell* **69**(2): 375-384.

Vaucheret, H., Vasquez, F., Crete, P. and Bartel, D.P. (2004). "The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development." *Genes Dev.* **18**(10): 1187-1197.

Vazquez, F., Vaucheret, H., Rajagopalan, R., Lepers, C., Gasciolli, V., Mallory, A. C., Hilbert, J. L., Bartel, D. P. and Crete, P. (2004). "Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs." *Mol Cell* **16**(1): 69-79.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S. I. and Moazed, D. (2004). "RNAi-mediated targeting of heterochromatin by the RITS complex." *Science* **303**(5658): 672-676.

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D. and 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**(10): 2508-2518.

Vogelauer, M., Wu, J., Suka, N. and Grunstein, M. (2000). "Global histone acetylation and deacetylation in yeast." *Nature* **408**(6811): 495-498.

Voit, R., Kuhn, A., Sander, E. E. and Grummt, I. (1995). "Activation of mammalian ribosomal gene transcription requires phosphorylation of the nucleolar transcription factor UBF." *Nucleic Acids Res* **23**(14): 2593-2599.

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. and Martienssen, R. A. (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." *Science* **297**(5588): 1833-1837.

Volpe, T., Schramke, V., Hamilton, G. L., White, S. A., Teng, G., Martienssen, R. A. and Allshire, R. C. (2003). "RNA interference is required for normal centromere function in fission yeast." *Chromosome Res* **11**(2): 137-146.

Vongs, A., Kakutani, T., Martienssen, R. A. and Richards, E. J. (1993). "Arabidopsis thaliana DNA methylation mutants." *Science* **260**(5116): 1926-1928.

Wallace, H., Birnstiel, M.L. (1966). "Ribosomal cistrons and the nucleolar organizer." *Biochim Biophys Acta.* **114**(2): 296–310.

Wallace, H., Langridge, W.H.R. (1971). "Differential amphiplasty and the control of ribosomal RNA synthesis." *Heredity* **27**: 1-13. (*cit in* Pikaard, 2000)

Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M. B., Rouse, D. T., Liu, Q., Gooding, P. S., Singh, S. P., Abbott, D., Stoutjesdijk, P. A., Robinson, S. P., Gleave, A. P., Green, A. G. and Waterhouse, P. M. (2001). "Construct design for efficient, effective and high-throughput gene silencing in plants." *Plant J* **27**(6): 581-590.

Workman, J. L. and Kingston, R. E. (1998). "Alteration of nucleosome structure as a mechanism of transcriptional regulation." *Annu Rev Biochem* **67**: 545-579.

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., Jacobsen, S. E. and Carrington, J. C. (2004). "Genetic and functional diversification of small RNA pathways in plants." *PLoS Biol* **2**(5): E104.

Yadegari, R., Kinoshita, T., Lotan, O., Cohen, G., Katz, A., Choi, Y., Nakashima, K., Harada, J. J., Goldberg, R. B., Fischer, R. L. and Ohad, N. (2000). "Mutations in the FIE and MEA genes that encode interacting polycomb proteins cause parent-of-origin effects on seed development by distinct mechanisms." *Plant Cell* **12**(12): 2367-2382.

Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., Pham, P., Cheuk, R., Karlin-Newmann, G., Liu, S. X., Lam, B., Sakano, H., Wu, T., Yu, G., Miranda, M., Quach, H. L., Tripp, M., Chang, C. H., Lee, J. M., Toriumi, M., Chan, M. M., Tang, C. C., Onodera, C. S., Deng, J. M., Akiyama, K., Ansari, Y., Arakawa, T., Banh, J., Banno, F., Bowser, L., Brooks, S., Carninci, P., Chao, Q., Choy, N., Enju, A., Goldsmith, A. D., Gurjal, M., Hansen, N. F., Hayashizaki, Y., Johnson-Hopson, C., Hsuan, V. W., Iida, K., Karnes, M., Khan, S., Koesema, E., Ishida, J., Jiang, P. X., Jones, T., Kawai, J., Kamiya, A., Meyers, C., Nakajima, M., Narusaka, M., Seki, M., Sakurai, T., Satou, M., Tamse, R., Vaysberg, M., Wallender, E. K., Wong, C., Yamamura, Y., Yuan, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R. (2003). "Empirical analysis of transcriptional activity in the Arabidopsis genome." *Science* **302**(5646): 842-846.

Yang, L., Huang, W., Wang, H., Cai, R., Xu, Y. and Huang, H. (2006). "Characterizations of a hypomorphic argonaute1 mutant reveal novel AGO1 functions in Arabidopsis lateral organ development." *Plant Mol Biol* **61**(1-2): 63-78.

Yekta, S., Shih, I. H. and Bartel, D. P. (2004). "MicroRNA-directed cleavage of HOXB8 mRNA." *Science* **304**(5670): 594-596.

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J. and Mello, C. C. (2006). "Analysis of the C. elegans Argonaute family reveals that distinct Argonautes act sequentially during RNAi." *Cell* **127**(4): 747-757.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V., Pellegrini, M., Goodrich, J. and Jacobsen, S. E. (2007). "Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis." *PLoS Biol* **5**(5): e129.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W., Chen, H., Henderson, I. R., Shinn, P., Pellegrini, M., Jacobsen, S. E. and Ecker, J. R. (2006). "Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis." *Cell* **126**(6): 1189-1201.

Zhang, Y. and Reinberg, D. (2001). "Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails." *Genes Dev* **15**(18): 2343-2360.

Zhao, X., Sternsdorf, T., Bolger, T. A., Evans, R. M. and Yao, T. P. (2005). "Regulation of MEF2 by histone deacetylase 4- and SIRT1 deacetylase-mediated lysine modifications." *Mol Cell Biol* **25**(19): 8456-8464.

Zheng, X., Zhu, J., Kapoor, A. and Zhu, J. K. (2007). "Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing." *Embo J* **26**(6): 1691-1701.

Zhou, Y. and Grummt, I. (2005). "The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing." *Curr Biol* **15**(15): 1434-1438.

Zilberman, D., Cao, X. and Jacobsen, S. E. (2003). "ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation." *Science* **299**(5607): 716-719.

Zilberman, D., Cao, X., Johansen, L. K., Xie, Z., Carrington, J. C. and Jacobsen, S. E. (2004). "Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats." *Curr Biol* **14**(13): 1214-1220.

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. and Henikoff, S. (2007). "Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription." *Nat Genet* **39**(1): 61-69.