CHAPTER TWENTY-ONE

Transcriptome Targets of the Exosome Complex in Plants

Dmitry Belostotsky

Contents

Abstract

The exosome complex is endowed with the capabilities to conduct 3′-end RNA processing, 3′-end degradation, and surveillance of various RNA substrates. Although the exosome is present in both eukaryotes and archaea, where it plays a central role in cellular RNA metabolism, a substantial degree of variability exists in its modus operandi in the different domains of life. This chapter discusses features of the exosome in plants that distinguish it from the exosome in archaea, fungi, and animals, as well as reviews the resources and tools that are needed to identify and catalog plant exosome targets on a transcriptome-wide scale.

1. Exosome: At the Nexus of the Cellular RNA Transactions

The exosome complex is an evolutionarily conserved macromolecular assembly with 3′ to 5′-exoribonucleolytic activity that is present in both nuclear and cytoplasmic cellular compartments. What the necessary brevity of this definition conceals is an overwhelming complexity and variability

Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri, Kansas City, Missouri, USA

Methods in Enzymology, Volume 448
ISSN 0076-6879, DOI: 10.1016/S0076-6879(08)02621-9

of exosome architecture and function. For example, the exosome can (1) process 3′-extended precursor RNAs to their mature products, (2) completely degrade other types of RNA substrates, and (3) selectively eliminate aberrant (e.g., misfolded) RNAs without affecting their normal counterparts. How substrate fate is determined in individual cases remains largely unknown. Moreover, which of the individual subunits of the exosome complex is responsible for its catalytic activity seems to vary across the phylogenetic spectrum, and depletion of the individual subunits from the exosome core engenders distinct sets of molecular phenotypes, depending on the species, tissue, and cell type. Hence, it is necessary to comprehensively define the RNA substrates that are processed or degraded by the exosome in various circumstances to fully grasp the impact of this macromolecular complex on gene expression, as well as to understand the reasons why its function is essential for viability.

The eukaryotic exosome core is an assembly of 9 or 10 polypeptides containing two main classes of subunits. One such class, represented by six polypeptides, has pronounced sequence similarity to *Escherichia coli* RNase PH (although these RNase PH-like polypeptides in fungi and animals lack detectable exoribonucleolytic activity). The second class contains proteins that possess S1- and KH-type RNA binding domains. In some, but not all, eukaryotic species, the exosome core also contains Dis3/Rrp44, a hydrolytic exoribonuclease. The crystal structures of the exosome from archaea revealed that the RNase PH type subunits are organized into a trimer of dimers, in a manner analogous to the prokaryotic phosphorolytic RNases (Lorentzen *et al.*, 2005). The archaeal exosome has three phosphorolytic exoribonuclease active sites within its core, although the width of its central channel seems too narrow to accommodate three RNA molecules simultaneously. Instead, the multiplicity of active sites may play a significant role in the processivity of the RNA degradation (Lorentzen *et al.*, 2007). Early ideas about the structural organization of the eukaryotic exosome likewise envisioned a heterohexameric ring of RNase PH-like subunits organized into a RNA processing chamber, with an RNA binding adaptor module containing the S1 and KH domain proteins, hereafter referred to as ''cap'', added on top of it (Aloy *et al.*, 2002). The modularity implicit in this model offered a potentially appealing solution to the problem of correctly choosing the substrate's fate whereby the events in the processing chamber would depend on the mode, of interaction of the RNA substrate and the adaptor module. Namely, depending on the stability of secondary structures in the RNA substrate and/or presence of RNA helicases in the accessory complexes (see below), the substrate would be either fully degraded or undergo a limited processing, followed by dissociation and release of the reaction product. However, as opposed to the apparent modularity of the exosome

complex in archaea, experimentally determined crystal structure of the reconstituted human exosome (Liu *et al.*, 2006 and chapter 10 by Greimann and Lima in this volume) shows that structural symmetry has considerably degenerated during evolution, such that every subunit in the human complex interacts with the other subunits through a unique interface. One corollary of this finding is that the integrity (and consequently function) of the eukaryotic exosome should strictly require the simultaneous presence of all subunits. Although this view is, indeed, consistent with results of genetic depletion experiments in yeast and *Trypanosama brucei*, it is not the case in plants, as discussed later.

A number of auxiliary factors interact with the exosome to facilitate its functions. The nuclear form of the exosome is distinguished by the presence of an additional subunit, RRP6, associated with the exosome core. The nuclear exosome is remarkably versatile and able to carry out either 3′-end-processing of RNA, such as in the case of the 5.8S rRNA precursor (Allmang *et al.*, 1999), or complete RNA degradation, as in the case of external transcribed rRNA spacers (Allmang *et al.*, 2000), aberrant pre-rRNAs, pre-mRNAs, pre-tRNAs (Bousquet-Antonelli *et al.*, 2000; Kadaba *et al.*, 2004; 2006; Libri *et al.*, 2002; Torchet *et al.*, 2002), or normal mRNAs trapped in the nucleus when mRNA export is blocked (Das *et al.*, 2003). Although the RNase D-like protein RRP6 is required for all activities of the nuclear exosome, a subset of these activities requires distinct auxiliary factors. For example, the putative RNA binding protein LRP1 participates in the processing of stable RNAs (Mitchell *et al.*, 2003; Peng *et al.*, 2003). On the other hand, the degradative activity of the nuclear exosome is linked to the *TR*F4/5-*A*IR1/2-*M*TR4 polyadenylation (TRAMP) complex, which helps recruit the exosome to aberrant or misfolded structural RNAs, cryptic unstable transcripts (CUTs), as well as to mRNAs that fail to complete splicing, 3′-end processing, or transport steps of mRNA biogenesis (Houseley and Tollervey, 2006; Houseley *et al.*, 2007; Kadaba *et al.*, 2004; 2006; LaCava *et al.*, 2005; Vanacova *et al.*, 2005; Wyers *et al.*, 2005).

In the cytoplasm, most of the activities of the exosome involve mRNA degradation (e.g., homeostatic mRNA turnover, rapid decay of unstable mRNAs, nonsense-mediated mRNA decay, degradation of mRNAs lacking stop codons and substrates of no-go mRNA decay [reviewed in Houseley *et al.*, 2006], and removal of 5′-products of RISC-mediated endonucleolytic mRNA cleavage) (Orban and Izaurralde, 2005). These reactions are mediated by the SKI2/SKI3/SKI8 complex and the SKI7 protein (Wang *et al.*, 2005). In addition, the exosome down regulates the levels of uncapped, unadenylated transcripts of LA virus, which is a cytoplasmic dsRNA virus present in many laboratory strains of yeast (Brown and Johnson, 2001). Hence, one major function of the cytoplasmic exosome is to destroy aberrant mRNAs and viral transcripts.

The 3′ to 5′-decay pathway mediated by the cytoplasmic exosome is a major mRNA decay pathway in mammals (Mukherjee *et al.*, 2002; Wang and Kiledjian, 2001). Specific RNA-binding proteins often play an important role in this process by acting as mediators of the exosome recruitment (e.g., as in the case of ARE [AU-rich element]) binding proteins (Chen *et al.*, 2001; Gherzi *et al.*, 2004). Interestingly, mammalian ZAP protein recruits the exosome to degrade viral mRNAs, thus suggesting that the exosome in mammalian cells also has an antiviral function (Guo *et al.*, 2007).

One important unresolved aspect of the coordination of mRNA degradation events in the cytoplasm concerns the spatial distribution of mRNA degradation reactions carried out by the exosome relative to those that are initiated by decapping. Although the decapping enzymes are concentrated in P bodies—distinct compartments for mRNA sequestration, storage, or decay (Sheth and Parker, 2003)—the exosome is not known to localize to P bodies (Brengues *et al.*, 2005). On the other hand, in the cultured Drosophila cells, some (but not all) of the exosome subunits are enriched in the cytoplasmic foci and, moreover, different subunits exhibit distinct localization patterns (Graham *et al.*, 2006). Perhaps this indicates a certain degree of plasticity in exosome composition and function. For example, one might envision that in higher cells there exist multiple kinds of specialized mRNA degradation subcompartments as opposed to just a single type in yeast. Therefore, different classes of mRNA may be routed to distinct sets of degradation sites, possibly depending on the cell type and/or physiological state. A more precise definition of spatially distinct exosome subcomplexes and a comprehensive global identification of their RNA substrates are required to address these issues.

2. Unique Features of the Plant Exosome

Although the ''parts list'' of the exosome is largely shared between plants and animals (and even archaea), the structure–function relationships in this macromolecular complex have been evolving independently for a long time. Indeed, the last common ancestor of plants and animals seems to have existed approximately 1.6 billion years ago (i.e., long before a clear fossil record of multicellular eukaryotes). Therefore, many informative lessons may result from comparative examination of the exosome composition, structure, localization, and RNA targets across the phylogenetic spectrum.

One major feature that distinguishes the plant exosome from its animal and fungal cousins is the presence of an active site in the RNase PH heterohexameric ring. Contrary to earlier proposals that were largely based on the extrapolations from structure–function studies of the archaeal

exosome complex (Aloy *et al.*, 2002), the RNase PH–like subunits of the yeast and mammalian complexes are catalytically inactive, and the hydrolytic subunit Dis3/Rrp44 is solely responsible for its activity (Dziembowski *et al.*, 2007). Concomitant with the addition of Dis3/Rrp44 to the complex in the course of evolution, exosomal RNase PH–type subunits have lost their catalytic competence through a divergence of amino acid residues responsible for the catalysis and/or binding to RNA and/or binding inorganic phosphate (which is essential for the catalytic mechanism of phosphorolytic enzymes). It has been hypothesized that RNase Dis3/Rrp44 acquisition may have decreased the selective pressure to maintain the phosphorolytic sites (Wahle, 2007). Yet, in contrast to the animal and yeast exosomes, all plants (and even the unicellular green alga *Chlamydomonas reinhardtii*), maintain the single phosphorolytic active site in the PH ring that resides in the RRP41 subunit (Chekanova *et al.*, 2000). In addition, no Dis3/Rrp44-like subunit copurified with the affinity-tagged *Arabidopsis* exosome (Chekanova *et al.*, 2007), which is also consistent with the above view.

Although the X-ray crystal structure of the human exosome (Liu *et al.*, 2006 and chapter 10 by Greimann and Lima in this volume) suggests that the structural integrity of the complex requires the simultaneous presence of every core subunit, a null allele of CSL4 (which is one of the RNA binding ''cap'' subunits in *Arabidopsis* has no apparent phenotype. Furthermore, although genetic depletion of any exosomal core subunit in *Saccharomyces cerevisiae* and *Trypanosoma brucei* results in virtually identical effects, at least on the RNA substrates examined so far, depletions of RRP4 and RRP41 in *Arabidopsis* produced overlapping yet distinct molecular signatures (and their null alleles led to different developmental phenotypes).

Taken together, these findings point to an unexpectedly high degree of functional plasticity in the plant exosome core, as well as call for broader examination of the structure–function relationships and targets of plant exosome complex. For example, it is important to explain the significance of the plant-specific conservation of the phosphorolytic active site in the RRP41. Moreover, the distinct effects resulting from the loss of the few tested subunits of the heterohexameric ''core'' and the RNA binding ''cap'' may reflect functional subfunctionalization of the complex into two submodules. This hypothesis can be tested by examining the phenotypes of additional representative core and cap subunits, as well as by studying the transcriptome-wide effects of their depletion. Finally, it remains to be established to what extent the unexpected features of the plant exosome can be generalized, particularly to monocot plants, which diverged ~230 million years ago from the dicot lineage, and of which *Arabidopsis* is a member. The tools and approaches necessary for such studies are discussed in the subsequent sections.

3. Resources for the Mutational Analyses of the Plant Exosome

The key prerequisite for inquiry into the function of the plant exosome is the availability of loss-of-function alleles. At present, the most commonly used access point for such alleles in *Arabidopsis* is the web site of the Salk Institute Genomic Analysis Laboratory (SIGnAL; signal.salk.edu/cgi-bin/tdnaexpress). The SIGnAL site represents an extensive, multifunctional web environment that not only provides convenient access to the SIGnAL's own sequence-indexed T-DNA insertional mutant library but also links to the multiple collections of *Arabidopsis* insertional mutants worldwide. The centerpiece of the SIGnAL's interface is a genome browser that contains links to a wide variety of relevant data for each of the annotated genome entries (e.g., transcriptome, methylome, small RNA), as well as a color-coded graphic display of the insertional events. Each insertion site symbol is a clickable link to information about which insertional mutagen was used, the flanking genomic sequence, and an interface for ordering seed stocks. A useful set of links to the individual papers and projects that more fully describe each set of tagged lines is provided at the bottom of the browser screen. As of late 2007, there were close to 400,000 tagged insertion lines publicly available. For a 119-Mb genome, this translates into a >96% chance of having a "hit" in every 1-Kb genomic segment. The single largest collection of the insertional mutants is the SIGnAL's own population of ~150,000 lines, produced with T-DNA insertions in the Columbia genetic background. In addition, SIGnAL has been generating a sub library of homozygous insertion lines for all nonessential genes, termed a "phenome-ready" genome set (i.e., ready for phenotyping). Although all core exosome subunit genes are essential in most species, CSL4 is dispensable in *Arabidopsis*, and this may hold true for some other core and noncore subunits. As of 2008, 17,637 homozygous insertion lines that represent 12,872 individual genes are available from ABRC.

Other sequence-indexed collections of *Arabidopsis* insertional mutants are available. The major ones include Syngenta (so-called SAIL lines) and GABI-Kat (Bielefeld University, Germany) T-DNA collections, the University of Wisconsin and Cold Spring Harbor Laboratory populations of stabilized *Ds* transposon insertions (all of the above, like the SIGnAL lines, are in the Columbia genetic background), FLAGdb T-DNA lines from INRA, France (in the WS background), as well as RIKEN (Japan) and the European EXOTIC consortium Ds element-based collections (in Nossen and Landsberg backgrounds, respectively). Information about the insertion sites for all of these resources can be accessed through SIGnAL. Although the distribution policies and procedures vary, most of mutant line seed stocks are distributed for a nominal fee by one of two major public

repositories'' - the *Arabidopsis* Biological Resource Center (ABRC) at the Ohio State University or the Nottingham Arabidopsis Stock Centre (NASC) at the University of Nottingham.

It is worth noting that SIGnAL's own collection has been created with a vector that contains the complete *35S* promoter from Cauliflower mosaic virus, including the enhancer. Therefore, insertion events may occasionally result in ''activation tagging'' instead of (or in addition to) the disruption of the normal expression of the nearby gene. Instances of unexpected and/or abnormal expression of flanking genomic regions (e.g., of antisense transcripts) have been reported (Ren *et al.*, 2004; and our unpublished observations). In addition, in the case of the SIGnAL lines, the kanamycin resistance gene that was used for the initial selection of transformants often gets silenced in subsequent generations and, hence, should not be used for selection or cosegregation analyses. This is less of a concern with hygromycin or BASTA resistance markers used in some of the other insertional collections. One final issue is that multiple T-DNA insertions per genome, as well as totally unrelated mutations (not linked to the T-DNA), are not uncommon and should be accounted for in interpreting the phenotypes.

Great strides have been made toward developing sequence-indexed collections of insertional mutants for other model plants. Among monocots, rice is the most advanced, where some 100,000 T-DNA insertion lines have been generated by Gynheung An's group in POSTECH (www.postech.ac.kr/life/pfg/risd; Jeong *et al.* [2002]). Additional T-DNA collections are available through Genoplante (France) and RIFGP (China), and Tos17 transposon insertion lines are available from NIAS (Japan). All of these are linked to the SIGnAL web site through the RiceGE (Rice Functional Genomic Mapping Tool) interface. The ~100,000 POSTECH lines are readily accessible by signing the MTA and provide an ~29.5% probability of a ''hit'' in any given rice gene on the basis of a genome size of 430 Mb, an average gene size of 3 kb, and 1.4 insertions per line (based on the relationship $P = 1 - [1 - f]^n$, where P is the probability of an insertion in any given gene, f is an inverse of the number of targets, and n is the total number of available T-DNA insertion events) (Clarke and Carbon, 1976).

A complementary resource is based on TILLING (*targeted induced local lesions in genomes*) technology. TILLING is a method of identifying EMS-induced point mutations in a gene of interest that relies on PCR amplification followed by the detection of heteroduplexes formed upon annealing of the mutant and wild-type PCR products using cleavage with a highly mismatch-specific endonuclease, CEL I (McCallum *et al.*, 2000). An Arabidopsis TILLING resource at Fred Hutchinson Cancer Center (tilling.fhcrc.org:9366) operates a high-throughput screening service on a fee-for-service basis. As of this year, the TILLING facility processed upward of 600 orders and reported nearly 8000 point mutations. Parallel TILLING resources have been established for rice (Till *et al.*, 2007), maize

(Till *et al.*, 2004), wheat (Slade *et al.*, 2005), and tomato (see tilling.ucdavis.edu/index.php/Main_Page). However, at present, TILLING is available as a public service only for maize (genome.purdue.edu/maizetilling) and rice (presently at the beta testing stage at the UC Davis Genome Center). Drawbacks of TILLING that should be kept in mind are its relatively high cost and the need for multiple outcrossing (because every genome experiences multiple "hits" during the typical mutagenesis procedure). Another approach to create mutants applicable to many plant species uses fast neutron mutagenesis, which tends to generate small deletions in the genome, coupled to a PCR-based method to identify abnormally small amplicons (Li *et al.*, 2001). It remains to be seen whether populations with a sufficient density of deletions can be generated to make this process efficient.

Another alternative approach involves the targeted engineering of a gene knockdown, which may be the only solution if an insertion or other type of mutation in the gene of interest is unavailable or not useful. For example, one could engineer an artificial microRNA (amiRNA) according to established thermodynamic constraints. This strategy has been successful in *Arabidopsis*, as well as rice (Ossowski *et al.*, 2008; Schwab *et al.*, 2006; Warthmann *et al.*, 2008), and it is particularly useful when one attempts to specifically target just one gene in a multigene family that consists of several closely related members. However, when discriminating among the members of a gene family is not of particular concern, higher efficiencies of gene knockdown can be achieved by creating transgenic plants that express relatively large (several hundred bp-long) double-stranded RNA hairpins, with the two arms of the hairpin separated by an intron or a nonspliced spacer (Smith *et al.*, 2000).

4. Transcriptome-wide Mapping of Targets of the Plant Exosome Complex

A caution is warranted regarding use of constitutive loss-of-function mutations to study the transcriptome-wide effects of exosomal subunits. Such mutations (or any constitutive mutations, for that matter) typically exhibit a complex mixture of primary and secondary effects (i.e. those that are caused by the chain of events whereby primary change in the expression of gene X in turn causes altered expression of gene Y etc) that are quite challenging to distinguish from one another, particularly on a transcriptome-wide scale. Usually, special considerations during the design stage and/or extensive follow-up efforts—both bioinformatic and experimental—are required to distinguish the primary effects from the secondary ones. This problem becomes particularly acute when deep sequencing or

tiling arrays are used for the analysis, because of the sheer volume of data generated by such approaches.

The most universal and radical approach that allows to bypass this issue is the engineering of a conditional allele, because in this case one can impose restrictive conditions at will and then monitor the real-time progression of the ensuing changes in the transcriptome, specifically focusing on the earliest time points of the response to the inactivation of the exosome. The expectation here is that the primary changes (e.g., those that directly result from the inactivation of the exosome complex) would develop early, whereas the secondary effects will become evident only after the significant alterations of the immediate substrates of the exosome accumulate (first in the mRNA, and subsequently in the protein levels). Another prediction would be that inactivation of a ribonucleolytic activity associated with the exosome would primarily result in an increase of immediate targets that are normally degraded by this complex, rather than their downregulation. On the other hand, those RNA species that normally are processed (rather than degraded) by the exosome would exhibit extensions beyond the point that normally correspond to the mature 3′-end of the molecule.

Conditional plant mutants in the core exosome subunit or its auxiliary factor can be constructed by means of an inducible expression of an artificial miRNA (Ossowski *et al.*, 2008; Schwab *et al.*, 2006), or by engineering an inducible RNAi (iRNAi) (Chekanova *et al.*, 2007; Chen *et al.*, 2003; Ohashi *et al.*, 2003; Smith *et al.*, 2000) to trigger a conditional knockdown of the desired target. One such iRNAi system successfully implemented in *Arabidopsis* is based on the estradiol-regulated chimeric transactivator XVE (Zuo *et al.*, 2000). To this end, transgenic *Arabidopsis* plants are engineered to express long (several hundred base pairs, if feasible) segments of the target gene as a pair of inverted repeats separated by an intron. Such intron-spliced hairpin RNAs tend to give higher silencing efficiency than those containing a nonspliced spacer between the arms of the hairpin (Stoutjesdijk *et al.*, 2002). Growing such exosome iRNAi plants on estradiol-containing media induces the RNAi-mediated knockdown of the respective targets and leads to growth arrest and subsequent death of seedlings. Importantly, a highly characteristic molecular phenotype characteristic of exosome malfunction (accumulation of the 3′-underprocessed 5.8S rRNA intermediates) was observed before any obvious effect on growth in both $RRP4^{iRNAi}$ and $RRP41^{iRNAi}$ lines (Chekanova *et al.*, 2007; Mitchell *et al.*, 1997). Such observations of characteristic molecular events that precede any overt effects on growth can be used to validate the transgenic iRNAi system. Hence, the corresponding time point of the depletion time course was used to harvest the plant tissues for extracting the RNA for the microarray experiments. In contrast, a similar analysis of the constitutive null mutant of CSL4 revealed massive amounts of secondary effects. This can seriously confound interpretations and limit the usefulness of the resulting data set (Chekanova *et al.*, 2007).

It should be noted that not all inducible RNAi systems are created equal, and hence a good deal of consideration must be given to their relative advantages, limitations, and potential associated artifacts. For example, iRNAi under the control of the popular glucocorticoid receptor/VP16/GAL4 DNA binding domain system (GVG) was found to trigger the expression of defense response genes on its own; similar observations have been reported by others in several plant species (Amirsadeghi *et al.*, 2007; Andersen *et al.*, 2003; Kang *et al.*, 1999). An excellent comprehensive review that evaluates numerous conditional promoter systems in multiple transgenic plants species has been recently published and should be consulted (Moore *et al.*, 2006).

The estradiol-regulated conditional RNAi knockdown lines were successfully used in conjunction with the whole genome tiling microarrays to comprehensively define the exosome targets in *Arabidopsis* (Chekanova *et al.*, 2007) and can be considered as a framework for future experiments (e.g., in other plants). The empty-vector control line and the iRNAi lines targeting the exosome components of choice are treated by estradiol for several days to deplete the respective exosome subunit. In our experience, germination for 5 days on half-strength MS media containing 8 μM of 17 β-estradiol (Sigma) results in approximately fivefold depletion in the level of the respective exosome subunit (measured by Q-PCR). Verification of the efficiency of the depletion is followed by the second RNA quality control to verify the 3′-under processing of the 5.8S rRNA by standard Northern blotting. RNA samples passing these quality control checks are then used to synthesize targets for the microarray hybridization. In addition to the estradiol-treated iRNAi lines and iRNAi lines exposed to just the solvent (DMSO) as a control (subsequently referred to as ± estradiol for brevity), the empty vector-transformed line should be processed similarly (hereafter referred to as WT for brevity). At present, most of the reagents and kits that are available off the shelf for *Arabidopsis* and other plants are designed for oligo(dT)-primed synthesis of the target. However, the exosome is believed to enter the pathway of the 3′ to 5′-mRNA decay only after the poly(A) tail is removed by one of several specialized deadenylating enzymes. Thus, many, if not most, of the mRNA decay intermediates accumulating as a result of the exosome depletion would be expected to lack the poly(A) sequence and, hence, would be missed in experiments with an oligo(dT)-primed target. This problem could be addressed by randomly priming the target instead of the use of the oligo(dT) priming, but doing so requires prior subtraction of the highly abundant ribosomal RNA. Although the reagents to do so are available for nonplant systems (e.g., Ribominus kit from Invitrogen), these are yet to be developed for subtracting plant ribosomal RNA.

The choice of an array platform for the studies of the plant exosome targets depends on a number of considerations (such as the availability of

commercial arrays for a given genome, equipment, core facilities, financial, and logistical constraints), but the whole genome tiling arrays are preferable because of a large number of noncoding RNAs regulated by the exosome (Chekanova *et al.*, 2007) that are largely missing from the annotation-based arrays. The Affymetrix whole-genome chip for *Arabidopsis* tiles both strands of its ~119 Mb genome end-to-end, with the average distance of 35 bases between the center points of the neighboring 25 base-long probes. Like most Affymetrix gene chips, each probe set consists of perfect match and mismatch probes (PM and MM), which presents the investigator with a choice of the use of MM signal as a measure of background probe binding. In practice, however, it has been observed that MM probes tend to still detect real signal in addition to nonspecific binding, hence subtracting the MM value from the PM signal value often adds to the noise instead of providing a gain in specificity (discussed in Irizarry *et al.* [2003]).

The transcriptome signals obtained by interrogating the tiling arrays with the exosome depleted line-derived targets can be analyzed with TileMap (Ji and Wong, 2005). The distinctive feature of the Tilemap algorithm is that it does not combine probes into groups to compute a summary metric for a given gene or a given region, but rather treats every probe as a separate entity. Hence, this software produces an unbiased identification of any genomic regions showing significantly altered expression level. The raw data are processed as follows.

The feature intensities for all replicates for the given knockdown line and the control line are quantile normalized and subjected to a multi-comparative TileMap analysis. For example, in the case of $rrp4^{RNAi}$, the data for a total of six tiling arrays (e.g., two replicates for WT treated with estradiol, two replicates for $rrp4^{RNAi}$ treated with estradiol, and two replicates for $rrp4^{RNAi}$ not treated with estradiol) from one strand were all quantile normalized at the same time. Differentially expressed regions are then identified by comparing the WT sample treated with estradiol and $rrp4^{RNAi}$ both with and without estradiol treatment, for both Watson and Crick strands. This generates two lists: (1) $rrp4^{RNAi}$ treated with estradiol >WT treated with estradiol as well as rrp4RNAi treated with estradiol > $rrp4^{RNAi}$ not treated with estradiol, and (2) $rrp4^{RNAi}$ treated with estradiol < WT treated with estradiol as well as $rrp4^{RNAi}$ treated with estradiol < $rrp4^{RNAi}$ not treated with estradiol. The posterior probability setting of 0.7, within a maximal region of 100 bases, can be used as a reasonable starting point in these analyses. The [$rrp4^{RNAi}$ treated with estradiol >$rrp4^{RNAi}$ not treated with estradiol] comparison identifies regions with increased expression because of the knockdown of Rrp4 protein levels, and the [rrp4^{RNAi} treated with estradiol >WT treated with estradiol] comparison removes regions with increased expression only because of estradiol alone.

To estimate the false-positive rate (i.e., transcripts that are induced or repressed by the estradiol treatment per se), replicates for WT sample with

and without estradiol treatment are quantile normalized and subjected to the TileMap analysis. Differentially expressed regions are identified by generating lists (WT with estradiol > WT no estradiol and WT treated with estradiol < WT with no estradiol) at posterior probability 0.7 within a maximal region of 100 bases (i.e., the same parameters as used for the comparisons of the knockdown lines). In our experience, this analysis yielded no differentially expressed regions. Moreover, even at a lower posterior probability value of 0.5, only six genomic regions could be identified (one upregulated and five downregulated). This is taken as evidence that any effects of estradiol alone can be considered negligible.

Application of the preceding tools in *Arabidopsis* has allowed the construction of the first high-resolution genome-wide atlas of the targets of the exosome complex in a multicellular eukaryote (Chekanova *et al.*, 2007). This analysis produced evidence for widespread polyadenylation-mediated and exosome-mediated quality control of plant RNA, as well as revealed the major categories of the plant exosome targets, including multiple classes of stable structural RNAs, a select subset of mRNAs, primary microRNA (pri-miRNA) processing intermediates, tandem repeat-associated siRNA precursor species, and numerous noncoding RNAs, many of which can only be revealed through repressing the exosome activity, because they remain below the level of detection in WT situation. This discovery of hundreds of noncoding RNAs that have not been previously described and hence belong to a ''deeply hidden'' layer of the transcriptome suggests that modulating the exosome activity, combined with further improvements in the sensitivity of detection beyond the one afforded by tiling microarrays, may help peel another layer off the transcriptome. This can be accomplished with deep sequencing technologies, such as Illumina 1G, ABI SOLiD, and/or 454 Life Sciences platforms.

ACKNOWLEDGMENTS

Research support from the grants from NIH, USDA, NSF, and BARD is gratefully acknowledged.

REFERENCES

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18,** 5399–5410.

Allmang, C., Mitchell, P., Petfalski, E., and Tollervey, D. (2000). Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res.* **28,** 1684–1691.

Aloy, P., Ciccarelli, F. D., Leutwein, C., Gavin, A. C., Superti-Furga, G., Bork, P., Bottcher, B., and Russell, R. B. (2002). A complex prediction: Three-dimensional model of the yeast exosome. *EMBO Rep.* **3,** 628–635.

Amirsadeghi, S., McDonald, A. E., and Vanlerberghe, G. C. (2007). A glucocorticoid-inducible gene expression system can cause growth defects in tobacco. *Planta* **226,** 453–463.

Andersen, S. U., Cvitanich, C., Hougaard, B. K., Roussis, A., Gronlund, M., Jensen, D. B., Frokjaer, L. A., and Jensen, E. O. (2003). The glucocorticoid-inducible GVG system causes severe growth defects in both root and shoot of the model legume *Lotus japonicus*. *Mol. Plant Microbe Interact.* **16,** 1069–1076.

Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell* **102,** 765–775.

Brengues, M., Teixeira, D., and Parker, R. (2005). Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* **310,** 486–489.

Brown, J. T., and Johnson, A. W. (2001). A cis-acting element known to block 3′-mRNA degradation enhances expression of polyA-minus mRNA in wild-type yeast cells and phenocopies a ski mutant. *RNA* **7,** 1566–1577.

Chekanova, J. A., Gregory, B. D., Reverdatto, S. V., Chen, H., Kumar, R., Hooker, T., Yazaki, J., Li, P., Skiba, N., Peng, Q., Alonso, J., Brukhin, V., *et al.* (2007). Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* **131,** 1340–1353.

Chekanova, J. A., Shaw, R. J., Wills, M. A., and Belostotsky, D. A. (2000). Poly(A) tail-dependent exonuclease AtRrp41p from Arabidopsis thaliana rescues 5.8 S rRNA processing and mRNA decay defects of the yeast ski6 mutant and is found in an exosome-sized complex in plant and yeast cells. *J. Biol. Chem.* **275,** 33158–33166.

Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Raijmakers, R., Pruijn, G. J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* **107,** 451–464.

Chen, S., Hofius, D., Sonnewald, U., and Bornke, F. (2003). Temporal and spatial control of gene silencing in transgenic plants by inducible expression of double-stranded RNA. *Plant J.* **36,** 731–740.

Clarke, L., and Carbon, J. (1976). A colony bank containing synthetic Col El hybrid plasmids representative of the entire *E. coli* genome. *Cell* **9,** 91–99.

Das, B., Butler, J. S., and Sherman, F. (2003). Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. *Mol. Cell. Biol.* **23,** 5502–5515.

Dziembowski, A., Lorentzen, E., Conti, E., and Seraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.* **14,** 15–22.

Gherzi, R., Lee, K. Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., and Chen, C. Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol. Cell* **14,** 571–583.

Graham, A. C., Kiss, D. L., and Andrulis, E. D. (2006). Differential distribution of exosome subunits at the nuclear lamina and in cytoplasmic foci. *Mol. Biol. Cell* **17,** 1399–1409.

Guo, X., Ma, J., Sun, J., and Gao, G. (2007). The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. *Proc. Natl. Acad. Sci. USA* **104,** 151–156.

Houseley, J., Kotovic, K., El Hage, A., and Tollervey, D. (2007). Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control. *EMBO J.* **26,** 4996–5006.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. *Nat. Rev. Mol. Cell. Biol.* **7,** 529–539.

Houseley, J., and Tollervey, D. (2006). Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO Rep.* **7,** 205–211.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., and Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4,** 249–264.

Jeong, D. H., An, S., Kang, H. G., Moon, S., Han, J. J., Park, S., Lee, H. S., An, K., and An, G. (2002). T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiol.* **130,** 1636–1644.

Ji, H., and Wong, W. H. (2005). TileMap: Create chromosomal map of tiling array hybridizations. *Bioinformatics* **21,** 3629–3636.

Kadaba, S., Krueger, A., Trice, T., Krecic, A. M., Hinnebusch, A. G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes Dev.* **18,** 1227–1240.

Kadaba, S., Wang, X., and Anderson, J. T. (2006). Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. *RNA* **12,** 508–521.

Kang, H. G., Fang, Y., and Singh, K. B. (1999). A glucocorticoid-inducible transcription system causes severe growth defects in *Arabidopsis* and induces defense-related genes. *Plant J.* **20,** 127–133.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121,** 713–724.

Li, X., Song, Y., Century, K., Straight, S., Ronald, P., Dong, X., Lassner, M., and Zhang, Y. (2001). A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant J.* **27,** 235–242.

Libri, D., Dower, K., Boulay, J., Thomsen, R., Rosbash, M., and Jensen, T. H. (2002). Interactions between mRNA export commitment, 3′-end quality control, and nuclear degradation. *Mol. Cell. Biol.* **22,** 8254–8266.

Liu, Q., Greimann, J. C., and Lima, C. D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127,** 1223–1237.

Lorentzen, E., Dziembowski, A., Lindner, D., Seraphin, B., and Conti, E. (2007). channelling by the archaeal exosome. *EMBO Rep* **8,** 470–476.

Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G., and Conti, E. (2005). The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat. Struct. Mol. Biol.* **12,** 575–581.

McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* **123,** 439–442.

Mitchell, P., Petfalski, E., Houalla, R., Podtelejnikov, A., Mann, M., and Tollervey, D. (2003). Rrp47p is an exosome-associated protein required for the 3′-processing of stable RNAs. *Mol. Cell. Biol.* **23,** 6982–6992.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: A conserved eukaryotic RNA processing complex containing multiple 3′ to 5′-exoribonucleases. *Cell* **91,** 457–466.

Moore, I., Samalova, M., and Kurup, S. (2006). Transactivated and chemically inducible gene expression in plants. *Plant J.* **45,** 651–683.

Mukherjee, D., Gao, M., O'Connor, J. P., Raijmakers, R., Pruijn, G., Lutz, C. S., and Wilusz, J. (2002). The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. *EMBO J.* **21,** 165–174.

Ohashi, Y., Oka, A., Rodrigues-Pousada, R., Possenti, M., Ruberti, I., Morelli, G., and Aoyama, T. (2003). Modulation of phospholipid signaling by GLABRA2 in root-hair pattern formation. *Science* **300,** 1427–1430.

Orban, T. I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. *RNA* **11,** 459–469.

Ossowski, S., Schwab, R., and Weigel, D. (2008). Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J.* **53,** 674–690.

Peng, W. T., Robinson, M. D., Mnaimneh, S., Krogan, N. J., Cagney, G., Morris, Q., Davierwala, A. P., Grigull, J., Yang, X., Zhang, W., *et al.* (2003). A panoramic view of yeast noncoding RNA processing. *Cell* **113,** 919–933.

Ren, S., Johnston, J. S., Shippen, D. E., and McKnight, T. D. (2004). Telomerase activator1 induces telomerase activity and potentiates responses to auxin in *Arabidopsis*. *Plant Cell* **16,** 2910–2922.

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

Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* **300,** 805–808.

Slade, A. J., Fuerstenberg, S. I., Loeffler, D., Steine, M. N., and Facciotti, D. (2005). A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat. Biotechnol.* **23,** 75–81.

Smith, N. A., Singh, S. P., Wang, M. B., Stoutjesdijk, P. A., Green, A. G., and Waterhouse, P. M. (2000). Total silencing by intron-spliced hairpin RNAs. *Nature* **407,** 319–320.

Stoutjesdijk, P. A., Singh, S. P., Liu, Q., Hurlstone, C. J., Waterhouse, P. A., and Green, A. G. (2002). hpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. *Plant Physiol.* **129,** 1723–1731.

Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S., and Comai, L. (2007). Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* **7,** 19.

Till, B. J., Reynolds, S. H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C. A., Enns, L. C., Odden, A. R., *et al.* (2004). Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* **4,** 12.

Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J., and Tollervey, D. (2002). Processing of 3′-extended read-through transcripts by the exosome can generate functional mRNAs. *Mol. Cell* **9,** 1285–1296.

Vanacova, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.* **3,** e189.

Wahle, E. (2007). Wrong PH for RNA degradation. *Nat. Struct. Mol. Biol.* **14,** 5–7.

Wang, L., Lewis, M. S., and Johnson, A. W. (2005). Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p. *RNA* **11,** 1291–1302.

Wang, Z., and Kiledjian, M. (2001). Functional link between the mammalian exosome and mRNA decapping. *Cell* **107,** 751–762.

Warthmann, N., Chen, H., Ossowski, S., Weigel, D., and Herve, P. (2008). Highly specific gene silencing by artificial miRNAs in rice. *PLoS ONE* **3,** e1829.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J. C., Dufour, M. E., Boulay, J, Regnault, B., Devaux, F., Namane, A., Seraphin, B., *et al.* (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121,** 725–737.

Zuo, J., Niu, Q. W., and Chua, N. H. (2000). Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.* **24,** 265–273.