Dicer-Dependent Biogenesis of Small RNAs and Evidence for MicroRNA-Like RNAs in the Penicillin Producing Fungus *Penicillium chrysogenum*

Tim A. Dahlmann, Ulrich Kück*

Christian Doppler Laboratory for “Fungal Biotechnology”, Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-Universität Bochum, Universitätsstr. 150, D-44780 Bochum, Germany

* ulrich.kueck@rub.de

Abstract

MicroRNAs (miRNAs) are non-coding small RNAs (sRNAs) that regulate gene expression in a wide range of eukaryotes. In this study, we analyzed regulatory sRNAs in *Penicillium chrysogenum*, the industrial producer of the β-lactam antibiotic penicillin. To identify sRNAs and microRNA-like RNAs (milRNAs) on a global approach, two sRNA sequencing libraries were constructed. One library was created with pooled total RNA, obtained from twelve differently grown cultures (RNA Mix), and the other with total RNA from a single submerged cultivation (Δku70FRT2). Illumina sequencing of both RNA libraries produced 84,322,825 mapped reads. To distinguish between Dicer-dependent and independent sRNA formation, we further constructed two single *dicer* gene mutants (Δdcl2 and Δdcl1) and a *dicer* double mutant (Δdcl2Δdcl1) and analyzed an sRNA library from the Dicer-deficient double-mutant. We identified 661 Dicer-dependent loci and in silico prediction revealed 34 milRNAs. Northern blot hybridization of two milRNAs provided evidence for mature milRNAs that are processed either in a complete or partial Dicer-dependent manner from an RNA precursor. Identified milRNAs share typical characteristics of previously discovered fungal milRNAs, like a strong preference for a 5’ uracil and the typical length distribution. The detection of potential milRNA target sites in the genome suggests that milRNAs might play a role in post-transcriptional gene regulation. Our data will further increase our knowledge of sRNA dependent gene regulation processes, which is an important prerequisite to develop more effective strategies for improving industrial fermentations with *P. chrysogenum*.

Introduction

Small regulatory RNAs are short non-coding RNAs ranging in size from 17 to 29 nt that mediate RNA interference (RNAi). Small RNAs (sRNAs) have been discovered in most eukaryotic organisms, and can be divided into three major regulatory classes [1–3]: piwi-interacting
RNAs (piwiRNAs), short interfering RNAs (siRNAs), and microRNAs (miRNAs). Their endogenous or extrinsic origins and different biogenesis pathways can distinguish these sRNAs from each other. SiRNAs and miRNAs are usually derived from double stranded RNAs (dsRNAs) that are processed by Dicer-like proteins, while piwiRNAs originate Dicer-independently from a single stranded RNA (ssRNA) [4,5]. In contrary to siRNAs, which are derived from double-stranded RNA molecules, miRNAs are processed from a nuclear-encoded single-stranded precursor RNA that forms a typical RNA stem-loop structure [6]. Besides exogenous siRNAs that mediate silencing of selfish genetic elements, endogenous siRNAs have been reported that regulate expression of endogenous genes in animals, plants, and fungi. These siRNAs originate from exonic regions (exonic-siRNAs), and target the mRNAs of protein coding genes from which they were produced [5,7–9].

Investigations into the Quelling phenomenon in *Neurospora crassa* led to the early discovery of regulatory sRNAs in a fungal organism [10]. This phenomenon is the silencing of multiple copies of homologous sequences, which were integrated by DNA transformation into fungal genomic DNA. In these cases, additional copies of the transgene induced the RNAi silencing machinery and siRNAs were processed by Dicer-like enzymes from dsRNAs generated by RNA-dependent RNA polymerases (RdRP) [11]. While the phenomenon of Quelling is well understood at the molecular level, little is known so far about fungal miRNAs. The first fungal microRNA-like RNAs (miRNAs) were discovered in *N. crassa* [12]. MiRNAs have also been detected in other fungal organisms, such as *Cryptococcus neoformans*, *Sclerotinia sclerotiorum*, *Trichoderma reesei*, *Fusarium oxysporum*, and *Penicillium marneffei* but their regulatory impact on gene expression and the mechanism of miRNA-target interaction is poorly understood [13–17]. In plants, near-perfect base pairing between the miRNA and its target is necessary, and results in direct cleavage of the target mRNA. In stark contrast, target recognition in animals requires a less stringent 7–8 nt long seed sequence at the 5’ end of the miRNA [18]. Besides directly degrading mRNA, animal and plant miRNAs can mediate suppression of mRNA translation [19,20]. Analysis of miRNAs from *N. crassa* and *C. neoformans* showed that miRNAs negatively regulate protein expressions of reporter genes, which contain perfect complementary target sites [12,13]. Nevertheless, the mechanisms involved in targeted gene silencing by fungal miRNAs are mostly unknown, and so far little information about target recognition is available.

Fungal miRNAs share many similarities with animal and plant miRNAs. All are similarly processed from stem-loop RNA precursors and the majority require a Dicer-like enzyme for their biogenesis [21]. In addition, Argonaute-like proteins, a core component of the RNA-induced silencing complex (RISC), are necessary to trigger gene silencing. Interestingly, the core components of RISC are highly conserved in diverse eukaryotes including fungi. With the exception of Saccharomycotina, the majority of ascomycetes analyzed so far contain at least one pair of Dicer-like and Argonaute-like protein coding genes [22,23].

Here, we analyzed sRNAs from *Penicillium chrysogenum*, the industrial producer of the β-lactam antibiotic penicillin. Previously, we showed that the genome contains two Dicer-like proteins [24]. Furthermore, a large (>200 nt) perfectly complementary artificial hairpin structure is processed by Dicer-like protein Dcl2 to generate functional siRNAs [24]. To detect miRNAs in *P. chrysogenum*, we analyzed different sRNA sequencing libraries from the industrial production strain P2niaD18, laboratory strain Δku70FRT2, lacking the Ku70 polypeptide of the non-homologous end joining (NHEJ) complex, and a Dicer-deficient mutant (Δdcl2Δdcl1) [25,26]. Sequencing of the small RNA transcriptome, followed by bioinformatic analyses to predict miRNAs and their putative targets, provide unequivocal evidence for Dicer-dependent sRNAs and miRNAs in *P. chrysogenum*.
Results

Analysis of small RNA libraries

For a comprehensive survey of sRNAs from *P. chrysogenum*, low molecular weight RNA (18–50 nt) was isolated and used for small RNA sequencing. For RNA isolation, the industrial laboratory strain P2niaD18 was grown on four different media and harvested at three different time points (see Material and Methods). From these 12 differently grown cultures, RNA was isolated and pooled to generate a sample designated "RNA-Mix". Two further small RNA libraries were generated from submerged cultures of Δku70FRT2 and Dicer-double mutant Δku70FRT2Δdcl2Δdcl1, hereafter referred to as Δdcl2Δdcl1. After removing low quality reads and trimming of 3’- adaptor sequences, 46,240,082, 52,144,415, and 65,660,252 trimmed reads were obtained as "total" read datasets for the RNA-Mix, Δku70FRT2 and Δdcl2Δdcl1. For subsequent analysis, the number of identical reads within each total dataset was calculated and all total reads were collapsed and transferred to a "unique" dataset. To avoid accumulation of rare reads, such as reads derived from degradation products, only unique reads with a read count of at least 10 were used for further analyses.

Total and filtered unique reads were mapped to the reference genome of P2niaD18 [25]. For the RNA-Mix, we found that 84.8% of all total reads (39,188,594) map at least once to the reference genome. Identical reads were combined to form a set of 1,117,770 unique reads and of these 122,018 were represented with a minimal read count of 10. The corresponding values of total reads for Δku70FRT2 and Δdcl2Δdcl1 that mapped at least once are 86.6% (45,134,231) and 87.9% (57,721,023) respectively. Within these samples we found 1,457,875 and 1,290,420 unique reads, respectively, including 113,603 and 104,922 unique reads with read numbers equal or higher than 10 (Table 1).

To determine the origin of the sRNAs, we mapped the sequence data acquired from the RNA-Mix, Δku70FRT2, and Δdcl2Δdcl1 to intergenic, exonic, and intronic regions and to rRNA and tRNA sequences. The proportion of reads that mapped to intragenic and intergenic regions, as well as tRNA and rRNA genes looked highly similar for all three libraries (S1 Fig) and is exemplarily displayed for Δku70FRT2 in Fig 1A and 1B. Within the total dataset, degraded rRNAs represented with 37.3% a rather high proportion. 33.9% of total reads mapped to intergenic regions and the amount of reads that mapped to introns (12.2%), exons (8.1%), and tRNAs (8.4%) was relatively low (Fig 1A). Interestingly, the number of total reads mapping forward to exonic sequences was lower than the number originating from the antisense sequences of these coding sequences (Table 1). Within the unique dataset of Δku70FRT2, the read distribution changed towards a higher fraction of reads mapping to exonic sequences (15.4%), while the percentage of reads from rRNA and tRNA regions declined to 30.2% and 3.6% (Fig 1B). To identify sRNA-producing loci, overlapping and adjacent unique reads were grouped. 1,693 loci were calculated for Δku70FRT2 and 783 for Δdcl2Δdcl1. Loci were mapped to the P2niaD18 genome and allocation of the loci to annotated genomic regions was performed. We calculated 1319 sRNA-producing loci in Δku70FRT2 and 374 of these generated sRNAs in forward and reverse orientation from the same genomic region. In Δdcl2Δdcl1, 783 sRNA-producing loci were identified but only 6 loci yielded sRNAs of both orientations and were located in intergenic regions or rRNA genes (Table 2). Pie graphs illustrate that the proportion of loci that map to different features has significantly been altered between Δku70FRT (Fig 1C) and the Dicer-deficient strain (Fig 1D). The percentage of sRNA-producing loci in exonic regions in sense orientation and the number of loci mapping to tRNAs and rRNAs genes have strongly increased in Δdcl2Δdcl1 compared to Δku70FRT. In particular, the significant depletion of sRNA-producing loci with sRNAs that map in both orientations is noteworthy and points to
the fact that Dicer-dependent sRNAs in *P. chrysogenum* are mostly produced from coding sequences and intergenic regions.

### Identification and characterization of Dicer-dependent small RNAs

Previously, we showed that a mutant lacking the gene coding for Dcl2 had a wild-type morphology, although siRNAs are not processed [24]. To identify Dicer-dependent sRNAs, read counts obtained from the Δku70FRT and Δdcl2Δdcl1 were normalized and compared. Furthermore, we characterized the identified sRNAs for specific traits of sRNAs, such as origin, length distribution, and starting nucleotide preference.

Length distributions of mapped reads ranging from 15 to 36 nt within the total and unique datasets are displayed in Fig 2A and 2B. Although the total reads were more or less irregularly distributed for all three samples, unique reads showed a significant peak at 21 nt for RNA-Mix and Δku70FRT2, but not for Δdcl2Δdcl1. In addition, we determined the first nucleotide of all unique reads. Recent reports have indicated a preference for regulatory sRNAs to start with uracil [12,27]. As expected for Dicer-processed sRNAs, unique reads obtained from RNA-Mix and Δku70FRT2 had an increased ratio of reads starting with uracil (58.0% and 49.2%). In Δdcl2Δdcl1 (20.9%), this phenomenon is not observed, obviously due to the absence of both Dicer-like proteins. To further demonstrate that the accumulation of reads starting with uracil depends on Dicer processing, all unique reads were calculated separately according to their read lengths. In Δku70FRT2, reads starting with uracil were predominantly found in the range between 18 to 23 nt, which is a common length of Dicer-processed sRNAs (Fig 2C) [12]. In contrary, no accumulation of reads with a 5'-uracil was present in Δdcl2Δdcl1 (Fig 2D).

### Table 1. Statistical summary of small RNA sequencing data and distribution of small RNAs with perfect match to the genome sequence of *P. chrysogenum*.

| RNAs (15–36 nt) | Total reads | Unique reads (≥ 10 reads) |
|-----------------|-------------|---------------------------|
|                 | RNA-Mix     | Δku70FRT2 | Δdcl2Δdcl1 | RNA-Mix | Δku70FRT2 | Δdcl2Δdcl1 |
| Raw reads       | 46,240,082  | 52,144,415 | 65,660,252 | 122,018 | 113,603 | 104,922 |
| Mapped reads[1] | 39,188,594  | 45,134,231 | 57,721,023 | 89,096  | 77,035  | 67,428  |
| Ratio           | 84.80%      | 86.60%    | 87.90%     | 73.01%  | 67.81%  | 64.26%  |
| Protein-coding genes |        |            |            |        |            |            |
| sense           | 11,612,935  | 10,975,391 | 13,497,944 | 29,659  | 27,835  | 24,798  |
| antisense       | 4,942,772   | 5,554,209  | 6,874,230  | 10,389  | 8,631   | 2,679   |
| Exon            | 5,619,395   | 6,472,049  | 7,539,001  | 16,527  | 17,680  | 8,431   |
| sense           | 1,633,831   | 1,687,351  | 1,557,211  | 9,788   | 11,875  | 7,596   |
| antisense       | 4,009,410   | 4,799,719  | 5,984,778  | 7,043   | 6,020   | 865     |
| Intron          | 10,569,963  | 9,777,092  | 12,504,834 | 21,702  | 17,439  | 18,022  |
| sense           | 9,996,408   | 9,276,048  | 11,918,300 | 19,744  | 15,839  | 17,085  |
| antisense       | 905,029     | 741,304    | 873,060    | 3,238   | 2,557   | 1,729   |
| tRNA            | 6,128,717   | 6,746,039  | 8,779,724  | 5,062   | 4,133   | 4,363   |
| sense           | 6,126,353   | 6,745,245  | 8,779,649  | 5,025   | 4,123   | 4,363   |
| antisense       | 2,577       | 971        | 277        | 38      | 11      | 1       |
| rRNA            | 22,659,643  | 29,792,848 | 39,201,820 | 43,143  | 34,808  | 39,639  |
| sense           | 22,652,991  | 29,780,783 | 39,201,243 | 43,051  | 34,639  | 39,626  |
| antisense       | 6,652       | 12,065     | 577        | 92      | 169     | 3       |
| Intergenic      | 21,367,882  | 27,127,757 | 35,564,101 | 50,226  | 41,195  | 35,646  |

[1] Perfectly mapped reads with more than one locus may occur in several categories as mapped reads.

doi:10.1371/journal.pone.0125989.t001
To identify Dicer-dependent sRNAs, all unique sRNAs in Δku70FRT2 were compared to those found in Δdcl2Δdcl1. After normalization of the datasets we identified 16,090 unique reads from Δku70FRT2 that were absent in the total read dataset of Δdcl2Δdcl1. Further, we found 3,695 unique reads that are strongly (log2 ≤ -2) and 4,734 that are slightly (log2 between -2 and -1) underrepresented in Δdcl2Δdcl1 compared to Δku70FRT2. On the contrary, the same analysis of the unique reads of Δdcl2Δdcl1 resulted in only 440 different reads that were not present in Δku70FRT2 (Fig 3A). The overwhelming majority of Dicer-dependent unique reads had a length between 17 and 23 nt and 64.8% start with uracil (Fig 3B), while Dicer-
independent sRNAs (Δdcl2Δdcl1) showed no significant preference for any nucleotide (S2 Fig). Mapping results of the 24,519 reads that were underrepresented by at least log2 \((-1\) in Δdcl2Δdcl1 compared to Δku70FRT2, were regarded as Dicer-dependent unique reads and revealed a different composition compared to the complete unique read dataset (Fig 1B and Fig 3C). The proportion of reads that mapped to exonic regions had increased (+13.7%), while the amount of rRNAs fragments had strongly decreased (-22.6%). The high percentage of reads originate from intergenic region did not show any significant change (-1.6%). These results led us conclude that the majority of Dicer-dependent sRNAs are processed from RNA molecules that are transcribed from intergenic, and exonic regions, especially from those in antisense orientation. Comparative analyses of the sRNA-producing loci from Δku70FRT2 and Δdcl2Δdcl1 showed that the accumulation of sRNAs in 661 sRNA-producing loci were affected by the Dicer-double deletion. In total, 368 loci that produced sRNAs in both orientations and 293 loci that produced sRNA in either sense or antisense orientation were found in Δku70FRT2, but were no longer present in Δdcl2Δdcl1. We further analyzed the distribution of these 661 Dicer-dependent sRNA loci. Within the 368 Dicer-dependent sRNA-producing loci that generated sense and antisense sRNAs, 246 were assigned to coding sequences and 121 to intergenic regions. The origin of the loci that showed a strand-specific accumulation of Dicer-specific sRNAs was determined, too. Of these 293 loci, 123 were assigned to exonic regions in sense and 35 in antisense orientation. Further, 131 were located in intergenic regions and 4 in tRNA genes, respectively (Fig 3D).

For representative genes, the distribution of sRNAs from Dicer-dependent and Dicer-independent loci on both strands are presented in Fig 4. Normalized read counts of the Dicer-dependent loci of the Copia13-like transposable element Pc17g00440 and the putative Helix-loop-helix DNA-binding protein Pc12g14660 are shown in Fig 4A. Besides Pc17g00440, multiple copies of the transposable element PCcopia13 (Pc17g00590, Pc21g00460, Pc22g26000, Pc24g01930, and Pc24g02680) and transposable element PCretro14 (Pc24g01940) were identified as Dicer-dependent loci with sRNAs in sense and antisense orientation. In Δdcl2Δdcl1, both strands showed a decreased level of sRNAs from transposable elements, while sRNAs from other exonic regions were only significantly decreased on one strand. A representative example for this is the gene Pc12g14660, coding a putative DNA-binding protein, that show a

Table 2. Characterization of small RNA producing loci.

| RNA-Mix       | Δku70FRT2 | Δdcl2Δdcl1 |
|---------------|-----------|------------|
| total number of sRNA-loci | 1513 | 1319 | 783 |
| loci with sRNAs on both strands | 438 | 374 | 6 |

| sRNA appearance | RNA-Mix | Δku70FRT2 | Δdcl2Δdcl1 |
|-----------------|---------|-----------|------------|
| Exon on both strands | 292 | 233 | 0 |
| only sense | 319 | 335 | 337 |
| only antisense | 137 | 114 | 20 |
| Intron on both strand | 2 | 1 | 0 |
| strand-specific | 21 | 17 | 16 |
| Intronic strand-specific | 177 | 181 | 176 |
| rRNA on both strands | 4 | 0 | 0 |
| strand-specific | 3 | 4 | 3 |
| rRNA strand-specific | 37 | 37 | 43 |
| Intergenic on both strands | 137 | 136 | 3 |
| strand-specific | 384 | 261 | 185 |

doi:10.1371/journal.pone.0125989.t002
significantly decreased sRNA level only on the antisense strand (Fig 4A). To illustrate sRNA distribution within Dicer-independent sRNA loci, putative cell-wall protein Pc20g06530 and a histidine-tRNA gene cluster are shown in Fig 4B. For these Dicer-independent loci, only sRNAs in sense orientation were detected. For Pc20g06530, sRNAs and mRNAs occur in similar amounts in both the recipient and double deletion strain Δdcl2Δdcl1. For all of the above mentioned examples, qRT-PCRs were performed to validate that a depleted amount of sRNA is not dependent from decreased transcription levels of mRNAs.

Identification of milRNAs in *P. chrysogenum*

The processing of microRNAs from an endogenous RNA hairpin molecule by Dicer-like enzymes is a major feature that differentiates milRNAs from other sRNAs. This feature allowed the *in silico* prediction of milRNAs by mapping sRNAs to afore predicted RNA-hairpin precursors. The position and frequency of sRNAs within the set of precursors was used to calculate milRNAs with miRDeep2 [28,29]. The RNA samples obtained from RNA-Mix and Δku70FRT2 contain Dicer-dependent sRNAs and were used for milRNA prediction, while

**Fig 2.** Length distribution of *P. chrysogenum* small RNA population. Length distribution of mapping sRNA reads for the datasets of total (A) and unique (B) reads obtained from three different samples. Frequency of the 5’-nucleotide of the unique reads of Δku70FRT2 (C) and Δdcl2Δdcl1 (D) in dependency of their read length.

doi:10.1371/journal.pone.0125989.g002
sRNAs from Δdcl2Δdcl1 were used as a negative control. For each dataset, milRNAs were predicted independently. 20 putative milRNAs were found for the RNA-Mix and 22 in Δku70FRT2. Eight milRNAs were found in both datasets, providing a total of 34 different milRNAs, designated milR-1 to -34. As expected for the Dicer-dependent formation of milRNAs, the number of corresponding reads were significantly reduced in Δdcl2Δdcl1. On the contrary, all predicted milRNAs found in Δku70FRT2 were also present in the RNA-Mix sample, and vice versa (Table 3). 31 of the 34 predicted milRNAs had a 5’ uracil, which is another indication for Dicer-processing of the predicted milRNAs, and the milRNA precursor sequences differed in length from 43 to 89 nt. The
Δdcl2 dataset helped to exclude predictions of milRNAs that were based on random patterns. Here, only two false-positives were found that did not show the typical features of predicted milRNAs, such as a 5’- uracil or the presence of a milRNA-star sequence.

All 34 milRNA precursors are located outside of tRNA and rRNA genes. The precursor sequence of most milRNAs occurred only once with a perfect match within the genome. Only the precursor sequence of milR-7 was found at three different loci (Table 3). To explore the location of milRNAs within the nuclear genome further, the mature sequence of milRNAs was
also mapped to the genome. Twenty-four mature miRNAs mapped only once to the reference genome, and 10 (miR-6, miR-7, miR-8, miR-11, miR-12, miR-13, miR-14, miR-26, miR-27, miR-30) mapped at least twice. The most frequent miRNA sequences are miR-13, and miR-11, with 13 and 8 loci. While miR-13 was found in exons, introns, 5' UTRs, and 3' UTRs, all loci of miR-11 were located within the 5' end of the coding sequence of multiple copies of a putative RNA-dependent DNA-polymerase (RdDP). This RdDP was previously identified as the Copia13-like transposable element PCcopia13 [30]. All copies of PCcopia13 have highly complementary loci, including the mature miR-11 and miR-11-star sequence.

| miRNA  | Sequence (5'-3') | Normalized Expression Level (TPTM[1]) | Score[2] | MFE[3] | Precursor location |
|--------|-----------------|----------------------------------------|----------|--------|-----------------|
| miR-1  | uugguaacgaugauuggaga | 122 | 305 | 1 | 882.2 | -41.0 | ChrI: 11548223–11548283, exon sense (Pc16g07970) |
| miR-2  | uggacgcaacgaugauuggaga | 152 | 3 | 0 | 373.4 | -17.1 | ChrII: 2410668–2410742, intergenic |
| miR-3  | uggacgcaacgaugauuggaga | 144 | 148 | 0 | 356.5 | -29.2 | ChrII: 1214806–1214889, exon antisense (Pc15g01560) |
| miR-4  | uggacgcaacgaugauuggaga | 6 | 40 | 0 | 208.3 | -22.4 | ChrII: 3088218–3088278, exon sense (Pc22g06050) |
| miR-5  | uggacgcaacgaugauuggaga | 17 | 27 | 0 | 231.2 | -24.2 | ChrIV: rev(37091–37159), exon antisense (Pc16g00090) |
| miR-6  | uggacgcaacgaugauuggaga | 69 | 123 | 0 | 167.2 | -32.7 | ChrI: 11697045–11697111, intergenic |
| miR-7  | uggacgcaacgaugauuggaga | 40 | 12 | 0 | 166.2 | -25.2 | ChrII: rev(11697681–11697758), intergenic; ChrI: rev(12068468–12068563), intergenic; ChrII: rev(4407094–4407171), intergenic |
| miR-8  | uggacgcaacgaugauuggaga | 12 | 5 | 0 | 128.5 | -28.7 |ChrI: 12126475–12126539, exon sense (Pc16g15350) |
| miR-9  | uggacgcaacgaugauuggaga | 23 | 2 | 0 | 135.4 | -35.1 | ChrIV: 2408756–2408823, exon antisense (Pc13g13430) |
| miR-10 | uggacgcaacgaugauuggaga | 66 | 14 | 0 | 90.4 | -30.0 | ChrI: rev(9237687–9237761), exon sense (Pc13g08790) |
| miR-11 | uggacgcaacgaugauuggaga | 19 | 1 | 0 | 75.6 | -26.9 |ChrII: rev(5605185–5605257), exon antisense (Pc21g03660) |
| miR-12 | uggacgcaacgaugauuggaga | 12 | 1 | 0 | 73.0 | -25.7 | ChrII: rev(8625875–8625949), exon antisense (Pc13g06180) |
| miR-13 | uggacgcaacgaugauuggaga | 52 | 27 | 1 | 104.6 | -24.4 |ChrII: rev(3729478–3729558), intergenic |
| miR-14 | uggacgcaacgaugauuggaga | 28 | 19 | 0 | 91.4 | -33.6 |ChrI: 7068824–7068895, intergenic |
| miR-15 | uggacgcaacgaugauuggaga | 66 | 14 | 0 | 90.4 | -30.0 | ChrI: rev(9237687–9237761), intergenic |
| miR-16 | uggacgcaacgaugauuggaga | 19 | 20 | 1 | 75.6 | -26.9 |ChrII: rev(3729478–3729558), intergenic |
| miR-17 | uggacgcaacgaugauuggaga | 12 | 1 | 0 | 73.0 | -25.7 | ChrII: rev(8625875–8625949), exon antisense (Pc13g06180) |
| miR-18 | uggacgcaacgaugauuggaga | 20 | 5 | 0 | 67.3 | -52.8 | ChrII: rev(3729478–3729558), intergenic |
| miR-19 | uggacgcaacgaugauuggaga | 10 | 1 | 0 | 61.9 | -29.1 |ChrII: rev(3729478–3729558), intergenic |
| miR-20 | uggacgcaacgaugauuggaga | 8 | 1 | 0 | 59.2 | -34.0 |ChrII: rev(3729478–3729558), intergenic |
| miR-21 | uggacgcaacgaugauuggaga | 15 | 2 | 0 | 39.9 | -32.8 |ChrII: rev(3729478–3729558), intergenic |
| miR-22 | uggacgcaacgaugauuggaga | 11 | 14 | 1 | 36.6 | -25.0 |ChrII: rev(3729478–3729558), intergenic |
| miR-23 | uggacgcaacgaugauuggaga | 10 | 4 | 0 | 36.6 | -21.7 |ChrII: rev(3729478–3729558), intergenic |
| miR-24 | uggacgcaacgaugauuggaga | 5 | 8 | 0 | 32.2 | -28.5 |ChrII: rev(3651114–3651202), intergenic |
| miR-25 | uggacgcaacgaugauuggaga | 8 | 6 | 0 | 31.6 | -19.8 |ChrII: rev(10506362–10506430), intergenic |
| miR-26 | uggacgcaacgaugauuggaga | 13 | 6 | 0 | 31.2 | -31.3 | ChrIII: rev(4059417–4059459), intergenic |
| miR-27 | uggacgcaacgaugauuggaga | 2 | 13 | 0 | 30.8 | -34.5 |ChrIII: rev(859369–859379), intergenic |
| miR-28 | uggacgcaacgaugauuggaga | 11 | 6 | 0 | 28.2 | -28.4 |ChrIII: rev(2351660–2351714), intergenic |
| miR-29 | uggacgcaacgaugauuggaga | 13 | 3 | 0 | 25.1 | -27.1 |ChrIII: rev(1906897–190926), exon antisense (Pc16g00780) |
| miR-30 | uggacgcaacgaugauuggaga | 29 | 12 | 0 | 24.0 | -24.5 | ChrIV: rev(3765153–3765197), intergenic |
| miR-31 | uggacgcaacgaugauuggaga | 2 | 4 | 0 | 22.6 | -36.1 | ChrIV: rev(719002–719055), exon antisense (Pc12g03150) |
| miR-32 | uggacgcaacgaugauuggaga | 5 | 1 | 1 | 22.4 | -31.6 | ChrIV: rev(1820276–1820326), intergenic |
| miR-33 | uggacgcaacgaugauuggaga | 15 | 1 | 0 | 20.8 | -19.0 | ChrIV: rev(1325461–1325507), intergenic |

[1] TPTM: transcripts per ten million
[2] Score: miRDeep2 prediction score
[3] MFE: minimal free energy

doi:10.1371/journal.pone.0125989.t003
MilRNA target gene prediction

So far, no algorithms are available that have been adapted to fungal milRNAs. Therefore, target predictions were performed using three different programs, written for either animals or plants. The program miRanda was written for accurately detecting microRNA binding sites in animals, and the programs TAPIR and psRNATarget were originally written for predictions in plants [31–33]. Since no 3'-UTR database is available for \textit{P. chrysogenum}, we extracted the 1 kb downstream sequences of all protein coding genes in the P2niaD18 genome. These 1 kb sequences were used as templates for target prediction. We succeeded in identifying mRNA targets for the predicted milRNAs described in this report. For 30 milRNAs, at least one target binding site was predicted with at least one program (S1 Table), and for 9 milRNAs (milR-6, milR-7, milR-8, milR-11, milR-14, milR-29, milR-30, milR31, milR-34) the same target sites were confirmed with programs written for the prediction of microRNAs in animals (miRanda) and plants (TAPIR and/or psRNATarget). Because some milRNAs are derived from 3'-UTRs, we compared the target sequence with the corresponding milRNA-star sequence, and only the predicted target sites of milR-8 and milR-14 are limited to their own milRNA-star sequences (Table 4). The largest number of putative target sites was predicted for milR-21; 118 binding sites of milR-21 were found within the 3'-UTRs of \textit{P. chrysogenum}. The most interesting mRNA targets were predicted for milR-1 and milR-21. For the most common milRNA, milR-1, the \textit{dcl2} (Pc12g13700) gene was predicted as target. For milR-21 at least three genes that might be involved in the DNA-repair mechanisms were predicted. These interesting findings led us to select milR-1 and milR-21 for further analyses.

Expression analysis and experimental validation of milRNAs

To confirm the presence of milRNAs in \textit{P. chrysogenum}, we performed northern blot analysis to detect milR-1 and milR-21. For northern blotting, total RNA from mycelia grown under the same conditions as for the library construction of \textit{Δku70 FRT2} and \textit{Δdcl2 Δdcl1} was used to detect predicted milRNAs. Both processed milRNAs as well as their precursor RNAs were detected by radioactive labeled probes (Fig 5). Interestingly, additional signals detected between the precursor and mature milR-21 may represent intermediates of the processed precursors. To test Dicer-dependent processing of milRNAs, further expression analyses were performed with total RNA from strains lacking one or both Dicer-like proteins. As expected for the \textit{dicer}-double deletion strain, milR-21 and its precursors were not detected. Furthermore, milR-21 and its precursors were also absent in the \textit{dcl2}-deletion, but not in the \textit{dcl1}-deletion strain, indicating that processing of milR-21 is Dcl2-dependent. Interestingly, weak signals for milR-1 and its precursors were detected in the Dicer-deficient strain \textit{Δdcl2 Δdcl1} and the \textit{Δdcl2} single deletion strain. These signals were significantly weaker than the signals observed from strains with an intact \textit{dcl2} gene. This result is consistent with the sRNA sequencing data presented in Table 3.

Discussion

Evidence for small RNAs in \textit{P. chrysogenum}

We previously reported that in \textit{P. chrysogenum} silencing of endogenous genes by artificial siRNAs is Dcl2 dependent [24]. Here we extend our analysis using small RNA deep sequencing of three different samples to increase our knowledge about endogenous fungal sRNAs, and specifically to identify microRNA-like RNAs. Indeed, we identified 34 milRNAs in \textit{P. chrysogenum} and demonstrate that the majority of sRNAs, including the novel milRNAs are processed from RNA molecules transcribed from intergenic and exonic regions, especially those in antisense
orientation. This agrees with a previously study in the ascomycete Fusarium oxysporum, where intergenic regions provided a high proportion of sRNAs [16]. In Magnaporthe oryzae it was further demonstrated that antisense sequences of coding regions can provide numerous sRNAs [34]. Further, a class of endogenous sRNAs that map to exons, called exonic-siRNAs, regulate the expression of protein coding genes in the basal fungus Mucor circinelloides [35]. To avoid analysis of mRNA degradation products, we used a strict cut-off for unique reads and identified sRNA-producing loci. We identified 661 Dicer-dependent sRNA-producing loci and demonstrate that 368 of these loci generate sRNAs in sense and antisense orientation from the same genomic source. The number of loci however that generate sRNA in both orientations is significantly decreased in the Dicer-deficient double-mutant. Further Dicer-dependent sRNAs are obtained mostly originating from intergenic and exonic regions in sense and antisense orientation. This observation is in accordance with the Dicer-dependent cleavage of dsRNAs and studies of exonic-sRNAs in M. circinelloides [35].

Although no obvious phenotype has so far been detected in \( \Delta dcl2 \Delta dcl1 \), we clearly demonstrated a lack of Dicer-dependent sRNAs in the Dicer-double mutant. Further, we showed that the amount of sRNAs on both strands of the selfish genetic element PCopia13 (Pc17g04400)

Table 4. miRNA target gene prediction.

| miR ID | target ID | function | miRanda score | TAPIR score | psRNA | UPE | Comment[1] |
|--------|-----------|----------|---------------|-------------|-------|-----|------------|
| miR-1  | Pc22g17450 | hypothetical protein | 164 | - | - | - | |
|        | Pc12g13700 | Dicer-like protein \((dcl2)\) | 161 | - | - | - | |
|        | Pc12g02820 | hypothetical protein | - | 4 | - | - | |
| miR-6  | Pc12g15260 | hypothetical protein | 174 | - | 17.2 | - | |
| miR-7  | Pc22g02960 | SNARE associated Golgi protein | 179 | - | 17.0 | - | |
| miR-8  | Pc04g00060 | hypothetical protein | 190 | 0 | 21.7 | star | |
|        | Pc19g00700 | hypothetical protein | 190 | 0 | 23.6 | star | |
|        | Pc24g01200 | hypothetical protein | 190 | 0 | 23.8 | star | |
|        | Pc24g02780 | hypothetical protein | 190 | 0 | 24.5 | star | |
| miR-11 | Pc16g15330 | hypothetical protein | 200 | 0 | 20.1 | star | |
|        | Pc12g06730 | hypothetical protein | 179 | 2.5 | - | - | |
| miR-14 | Pc24g01250 | hypothetical protein | 195 | 0 | 13.3 | star | |
|        | Pc17g00230 | hypothetical protein | - | 1 | 14.4 | - | |
|        | Pc21g00150 | putative proline-rich call wall protein | - | 2 | 15.2 | - | |
|        | Pc24g02850 | hypothetical protein | - | 2 | 17.1 | - | |
| miR-21 | Pc21g17740 | SNF2 family helicase | 290 | - | - | - | |
|        | Pc12g12880 | ubiquitin carboxyl-terminal hydrolase | 286 | - | - | - | |
|        | Pc13g12300 | BRCA1-like DNA repair protein | 283 | - | - | - | |
|        | Pc13g05360 | putative Ku70-binding protein | - | - | 20.8 | - | |
| miR-29 | Pc20g05760 | putative cytochrome P450 protein | 190 | 0 | 18.5 | star | |
|        | Pc18g04580 | hypothetical protein | 170 | 3 | 19.9 | - | |
| miR-30 | Pc22g16980 | hypothetical protein | 169 | 3.5 | 21.2 | - | |
| miR-31 | Pc22g13440 | hypothetical protein | 185 | 0 | - | star | |
|        | Pc21g05840 | putative Zn(2)-C6 DNA-binding protein | 176 | 2.5 | - | - | |
|        | Pc21g02000 | PLC-like phosphodiesterase | 171 | 4 | - | - | |
|        | Pc12g10520 | hypothetical protein | 172 | 3.5 | - | - | |
| miR-34 | Pc16g04680 | hypothetical protein | 195 | 0 | 19.9 | star | |
|        | Pc18g05650 | hypothetical protein | 171 | - | 21.7 | - | |

[1] Predicted target sites were compared to the miRNA-star sequence.

doi:10.1371/journal.pone.0125989.t004
was significantly reduced in the Dicer-deficient double-mutant. This observation is consistent to results obtained for transposons from *Neurospora crassa* and *Magnaporthe oryzae* and for the antiviral defense mechanism in the chestnut blight fungus *Cryphonectria parasitica* [36–38]. Genetic stability is an essential feature for industrial applications using *P. chrysogenum* [39]. Therefore, a well-functioning defense mechanism against selfish transposable elements is an important requirement of industrial strains to guarantee stable production yields.

Further, we observed typically characteristics for sRNAs known from other filamentous fungi, like significant preference for a 5'-uracil, their length distribution, and their origin from intergenic and exonic regions [40]. However, the percentage of uracil-starting Dicer-dependent sRNAs was only drastically higher for sRNAs ranging from 17 to 23 nt. Nevertheless, Dicer-dependent sRNAs were the predominant class of sRNAs in the overall sRNA population. We have demonstrated that formation of milR-1 is only partially Dicer-dependent. This result infers that at least one Dicer-independent pathway may be present to generate milR-1 in the Dicer-deficient mutant. This idea is strengthened by the presence of various Dicer-independent pathways of interfering sRNA in *N. crassa* [12]. Therefore, further analyses are necessary to reveal Dicer-independent sRNA-biogenesis pathways in *P. chrysogenum*.

Fig 5. Validation and expression analysis of (A) milR-1 and (B) milR-21. Total RNA from strain P2niaD18, the recipient Δku70FRT2 as well as Dicer single and double mutant strains were used for polyacrylamide gel electrophoresis and northern blot analysis. Mature milRNAs (milRs) and milRNA precursors (pre-milRs) were detected with reverse complement 32P-labeled DNA probes. Below, loading controls of the total RNA, stained with ethidium bromide (EtBr), and the predicted secondary structures of milRNA precursors are given. On the secondary structures, milRNA sequences are highlighted in red and arrows indicate the expected Dicer cleavage sites.
Evidence for miRNAs

Our bioinformatics analysis identified 34 Dicer-dependent sRNAs that carry the typical features of conventional miRNAs as similarly characterized in other eukaryotes. Typical features are RNA stem-loop formation, a strong preference for uracil at the 5’ end, and an average size of about 21 nt [5,41]. We predicted 8 milRNAs in the small RNA sequencing datasets of both RNA-Mix and Δku70FRT2, and repeatedly deriving these 8 milRNAs demonstrates the reliability of our analysis. Due to different growing conditions, some milRNAs were only predicted in the RNA-Mix or Δku70FRT2 dataset. Nevertheless, all mature milRNAs were represented by reads within both datasets. This indicates different expression rates of milRNA under different growth conditions. Further, these reads were either totally absent or reduced to single reads in the Δdcl2Δdcl1 double mutant.

Since knowledge about fungal milRNAs is still limited, improving parameters to predict fungal milRNAs is a challenging task for the future. However, we were able to identify Dcl2-dependent milRNAs ranging in size between 19 and 29 nt, which is consistent with previous findings that some mature fungal milRNAs are longer than regular animal and plant miRNAs [12,16]. Nevertheless, the majority of fungal Dicer-dependent sRNAs and milRNAs range in size between 19 and 22 nt. All predicted mature milRNA sequences of *P. chrysogenum* were compared to those recently found in other fungi, but no conserved milRNAs were found.

The predicted precursor sequences range in size between 43 to 89 nt. These results are consistent with milRNAs and milRNA-precursors found for other ascomycetes [12,14–16]. Northern blot analyses confirmed the existence of precursor and mature milRNAs milR-1 and milR-21 in *P. chrysogenum*. So far, we were unable to detect pri-milRNAs precursors, most probably due to their low abundance. When analyzed in detail, both milRNAs proved to be complete or partial Dcl2 dependent; moreover Dcl1 and Dcl2 did not show functional redundancy. These observations are also supported by our previous analyses demonstrating that only Dcl2 is essential for processing artificial dsRNAs to generate functional siRNAs that silence genes containing a perfect complementary target site [24]. These results are consistent with findings of milRNA biogenesis in *P. marneffei* but in contrast to findings in *N. crassa*, where both Dicer-like proteins are functionally redundant [12,17,42].

Remarkably, milR-1 was detected in the Dicer double-deletion strain. This demonstrates that milR-1 can form Dicer-independently, and that an alternative Dicer-independent pathway may exist to generate milRNAs. In *N. crassa* at least four different milRNA-generating pathways have been discovered involving components of Dicer-like proteins, Argonaute proteins, exonucleases (QIP) and an RNase III domain-containing protein (MRPL3). It was also demonstrated that single milRNAs can be processed independently by two different pathways [12].

Here, we clearly demonstrate that in *P. chrysogenum* the majority of sRNAs are Dicer-dependent and that Dcl2 is essential for the maturation of milR-21 and necessary for the accumulation of milR-1. Since the majority of predicted milRNAs are not present in Dicer double-deletion strain Δdcl2Δdcl1, their generation is Dicer dependent.

Among the predicted 34 milRNAs, 24 map only once to the reference genome, and 10 at least twice. Surprisingly, one of the most abundant milRNA sequences was milR-11, found 8 times in the 5’ end of the coding sequence for the Copia13-like transposable element PCopia13. We demonstrated that the level of sRNAs from this transposable element was reduced on both strands in Δdcl2Δdcl1. This indicates strongly to an sRNA-mediated RNAi pathway of selfish genetic elements in *P. chrysogenum*. The location of milRNA genes inside transposable elements is consistent with recent studies in *Cryptococcus neoformans*, where multiple loci of two milRNAs were discovered in the sequences of transposable elements and pseudogenes. It was demonstrated that milRNAs in *C. neoformans* induce transgene silencing via the canonical
RNAi pathway [13]. Our results prompt us to believe that miRNAs in *P. chrysogenum* are likewise involved in the negative regulation of transposon activity.

**Various mRNA targeting by milRNA?**

Using diverse programs, we were able to predicted potential milRNA target sites for all 34 miRNAs on putative 3’-UTRs of *P. chrysogenum*. Eight milRNA binding sites, calculated by miRanda, were validated by the plant miRNA-prediction programs TAPIR and/or psRNATarget. While some miRNAs had only a single predicted target, others, like milR-21, have the option of interacting with multiple mRNA targets. Interestingly the second best predicted target for milR-1 is the *dcl2* gene (Pc12g13700) and milR-21 was predicted to bind to at least three mRNA targets that are involved in DNA-repair mechanisms. As yet, we were not able to detect miRNA-mediated mRNA-degradation in *P. chrysogenum*. Previously, reporter gene constructs with artificial target sites were used to demonstrate milRNA-mediated silencing in fungi [12,13]. In these studies, target-mRNA degradation has only been demonstrated with reporter genes that carried a fully complementary target site. In contrast to siRNAs, miRNAs do not require full complementarity to bind and mediate silencing of target mRNAs. Thus, it is suggested that miRNA-mediated gene silencing in fungi is mostly caused by translational repression like it is known for animals and plants [1,12,43]. Therefore, further analyses of target-mRNA protein expression are necessary provide evidence for miRNA-target interactions and thus, miRNA mediated gene silencing in *P. chrysogenum*. Nevertheless, our small RNA sequencing approach identified novel miRNAs that are predicted to bind multiple target mRNAs, possibly as factors involved in a variety of different biological processes, like DNA-repair. Finally, results from sRNA identification and milRNA target predictions will serve as valuable clues for understanding sRNAs as regulators of gene expression in *P. chrysogenum*. A deeper understanding of sRNA-mediated gene regulation is an important prerequisite for developing more effective strategies to genetically manipulate this industrially important fungus to further improve production processes.

**Materials and Methods**

**Strains and growth conditions**

For the RNA-Mix, *P. chrysogenum* strain P2niaD18 was grown as submerged cultures in 500 ml shaking flasks in 100 ml CCM and MM at 27°C and 120 r.p.m., and as surface cultures on PAL Science Supor 200 Disc Filters on solid Oatmeal and M322 media. All cultures were inoculated with 1.0 x 10⁷ conidiospores as previously described [44,45]. Liquid cultures were grown for 72, 96, and 120 hours, and surface cultures for 48, 72, and 96 hours. The mycelia of these 12 different cultures were frozen in liquid nitrogen and total RNA was isolated separately as described in the "Nucleic acid isolation" section. After purification, total RNA of the 12 samples were pooled in equal amounts to the so-called RNA-Mix sample. For the samples Δku70FRT2 and Δdcl2Δdcl1 the recipient strain Δku70FRT2 and the Dicer-deficient double-mutant Δdcl2Δdcl1 were grown for three days in shaking flasks containing CCM under the same conditions as mentioned above (S2 Table and S3 Table).

**Identification of RISC core components**

The recently published genome of *P. chrysogenum* strain P2niaD18 [25] and the previously published annotation of strain Wisconsin 54–1255 [46] served as source for identifying genes involved in RNAi. To identify core components, BLAST analyses of RISC proteins from *N. crassa* were performed (S3 Fig and S4 Table) [47]. Protein sequence data from *N. crassa* were
obtained from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/), and sequences were aligned using MAFFT [48]. For domain identification, the predicted *P. chrysogenum* protein sequences were used to perform an NCBI conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/) [49].

**Generation of Dicer single- and double-knockout strains**

Plasmids were constructed to substitute genes encoding Dicer-like proteins in *P. chrysogenum* recipient strain Δku70FRT2 (S5 Table). For homologous recombination, fragments of about 1 kb flanking the 5' and 3' ends of the coding sequence were amplified by PCR and ligated to corresponding 5' and 3' flanks of nourseothricin or phleomycin knockout cassettes, expressing the nat or ble gene, under control of the trpC promoter. For deletion of dcl2 (Pc12g13700), the gene was replaced with a phleomycin resistance cassette. The nat-knockout cassette, used for deletion of dcl1 (Pc21g06890), carries additional components of the FLP/FRT recycling system that are under the control of the inducible promoter xylP [50]. Transformation of *P. chrysogenum* strains with foreign DNA was performed using conventional transformation procedures with modifications as previously described [45,51]. Homologous integration of the knockout construct and substitution of the target genes were confirmed by Southern blot and PCR analysis covering regions within and flanking the knockout cassette (S4 Fig, S5 Fig and S6 Fig). With this procedure, we generated dcl1- and dcl2-deletion strains. Transformation of plasmid pKODcl1 in dcl2-deletion strain resulted finally in the construction of the Δdcl2Δdcl1 double-deletion mutant.

**Nucleic acid isolation**

As described in the "Strains and growth conditions" sections, total RNA used for small RNA sequencing was obtained from different cultivations. For the "RNA-Mix" sample, total RNA was isolated independently from twelve different cultures of strain P2niaD18. For samples Δku70FRT2 and Δdcl2Δdcl1 a single culture of the corresponding strains was grown in liquid CCM for 72 h. For all samples, mycelia were collected for guanidinium-thiocyanate-phenol-chloroform extraction of total RNA [52]. RNA was measured with NanoDrop 1000 Spectrophotometer and Bioanalyzer 2100, using the Agilent RNA 6000 Nano Kit. All total RNAs that were pooled to equal amounts for the RNA-Mix show RNA integrity numbers (RIN) ≥ 9.0. For the analysis of Δku70FRT2 and Δdcl2Δdcl1 only total RNA with RIN values ≥ 9.8 were used. The 28S:18S ratio was calculated with 1.7 for all samples. DNA preparation was carried out as described recently [53].

**Small RNA deep sequencing**

Total RNAs from the three different samples described above in the "Strains and growth conditions" section were used for small RNA library preparation and small RNA sequencing. Library preparation and sequencing was performed at GATC Biotech, Konstanz, Germany. Total RNA samples were separated by a 10% TBE-urea denaturing polyacrylamide gel electrophoresis (PAGE) and sRNAs ranging between 18 and 50 nt were used for library preparation using the Illumina TruSeq Small RNA Sample Preparation Kit. After reverse transcription and amplification, cDNA products were checked and measured with Bioanalyzer 2100. Sequencing was performed on an Illumina HiSeq 2000 platform. Trimming of TruSeq adapter sequences was performed with the program Cutadapt and only reads trimmed at their 3’ end were used for further studies [54]. Raw sequencing data from sRNA sequencing have been deposited in the NCBI sequence read archive (SRA) (http://www.ncbi.nlm.nih.gov/sra) under the accession numbers SRR1705825 (RNA-Mix), SRR1706009 (Δku70FRT2), and SRR1706010 (Δdcl2Δdcl1).
Small RNA data analysis and miRNA prediction

To examine the origin of the cleaned RNA reads, alignments against the reference genome of *P. chrysogenum* P2niaD18 and diverse datasets, containing all intronic, exonic, and intergenic regions, as well as tRNA and rRNA sequences were performed with bowtie v1.1.0 and only perfect matches were considered [25,55]. This procedure was performed for total reads and unique reads with a read count of at least 10 that were collapsed using the perl script mappler.pl, a part of miRDeep2 [29]. For the calculation of read length and starting nucleotide distribution custom perl scripts were written. To remove spurious sRNAs, derived by RNA degradation, we adapted a previously published method [56]. All unique overlapping or closely mapped sRNAs were grouped. Only loci with at least four unique reads within a sliding interval of 300 bp were taken into account for further analyses. To identify Dicer-dependent reads and loci, unique reads and loci obtained from Δku70FRT2 and Δdcl2Δdcl1 were compared to each other. Read counts were normalized (transcripts per ten million, TPTM) using all reads that map to intragenic and intergenic regions. Reads mapping to structural RNAs, like tRNAs or rRNAs, were not considered for normalization. After normalization, log2 ratios were calculated using the count number of all unique reads obtained from the Dicer-deficient mutant and Δku70FRT2. To distinguish between Dicer-dependent and-independent loci, all loci in Δku70FRT2 and Δdcl2Δdcl1 were compared in accordance to their strand bias. A genomic region that showed sRNAs either on one strand or in forward and reverse orientation was rated as a sRNA-producing locus. Loci that appeared in Δku70FRT2 but not in the Dicer-deficient double-mutant Δdcl2Δdcl1 were assigned as Dicer-dependent.

MilRNAs were predicted with the program miRDeep2 with a score cutoff of 20 and without using additional miRNA information. For all predicted milRNAs, the read counts inside all three unique datasets were determined. For visualization of predicted milRNA precursors, RNAfold, part of the Vienna RNA Package 2.0, was used to predict secondary structures [57].

MilRNA target prediction

Little is known about milRNA-target interactions in fungi. To include both animal and plant surveys into our analysis, the program miRanda, written to predict miRNA targets in animals, and the programs TAPIR and psRNATarget, to predict potential plant-like target interactions, were used with default settings [31–33]. Since no 3’-UTR database is available for *P. chrysogenum* the 1 kb downstream sequences were extracted from all protein coding genes. These hypothetical 3’-UTRs were used for milRNA target prediction.

Northern blot analyses

For northern blot analysis, conducted to confirm the existence of predicted milRNAs, 30 μg of total RNAs were separated on a 15% denaturing polyacrylamide gel, containing 7 M urea. Separated RNAs were transferred onto a Hybond-N+ hybridization membrane (GE Healthcare Life Sciences) and crosslinked by UV-radiation (Stratalinker). Synthesized DNA StarFire probes (DNA Integrated Technology), which were perfectly complementary to the mature milRNA were labeled with α-[32P] dATP, following the manufacturer’s instructions. Hybridization in modified Church-Gilbert-hybridization solution (0.25 M KH2PO4, 0.25 M K2HPO4, 7% SDS, and 1 mM EDTA) was started at 65°C for 1 h, followed by an overnight stepwise cooling phase ending at 35°C [58,59]. For detection of bound radioisotopes Fujifilm BAS-III PhosphorImager plates and FLA-3000 detection system were used.
Quantitative real-time PCR

Quantitative real-time PCR (qRT-PCR) was carried out in a StepOnePlus real-time PCR system (Applied Biosystems) with GoTaq qPCR MasterMix for SybrGreen (Promega) as described previously [60,61]. Oligonucleotide primers are given in S6 Table.

Supporting Information

S1 Fig. Annotation of small RNA loci of all three datasets, RNA-Mix, Δku70FRT2, and Δdcl2Δdcl1. Pie graphs show the proportion of small RNAs mapped to intergenic, exonic and intronic regions or rRNA and tRNA genes, for total reads and unique reads. (TIF)

S2 Fig. Nucleotide preference of Dicer-dependent and-independent small RNAs. Beside the strong preference for uracil at the 5'-end for Dicer-dependent reads, no further nucleotide preference was detected for other positions inside Dicer-dependent and-independent sRNAs. (TIF)

S3 Fig. Comparative analysis of the Dicer-protein domains of DCL-2 and DCL-1 in N. crassa and Dcl2 (Pc12g13700) and Dcl1 (Pc21g06890) in P. chrysogenum. Conserved domains within the proteins are indicated and alignments are displayed below. (TIF)

S4 Fig. Construction and validation of Δdcl2 mutants. Replacement of the gene coding for Dcl2 (Pc12g013700), with phleomycin resistance cassette containing the Streptoalloteichus hindustanus (Sh) ble gene. Validation of homologous integration within the recipient strain Δku70FRT2 was performed by Southern blotting of 20 μg EcoRV digested genomic DNA with 32P-labeled complementary DNA probes of the 5'- (dark bar) and 3'-flank (bright bar) of dcl2. GeneRuler DNA Ladder (Thermo Scientific) was used as size standard. Homokaryotic transformants that show the expected fragments according to a correct genomic integration of the resistance cassette are marked with an asterisk. (TIF)

S5 Fig. Construction and validation of Δdcl1 mutants. Replacement of the gene coding for Dcl1 (Pc21g06890), with the inducible FLP/FRT cassette containing an N-acetyltransferase coding gene (nat) that mediates resistance to the antibiotic nourseothricin. Validation of homologous integration within the recipient strain Δku70FRT2 was performed by Southern blotting of 20 μg BglII digested genomic DNA with 32P-labeled complementary DNA probes of the 5'-flank (dark bar) of dcl1. Furthermore, PCR analyses, with primers (indicated with grey arrows) surrounding the sequences used for homologous integration, validate the homologous integration of the resistance cassette. In contrast to the tested transformants, the no-template controls (-Control) and PCRs with the recipient DNA (WT) show no PCR product. GeneRuler DNA Ladder (Thermo Scientific) was used as size standard. Homokaryotic transformants that show correct genomic integrations are marked with an asterisk. (TIF)

S6 Fig. Construction of Δdcl2Δdcl1 double mutants based on the strain Δdcl2. Homologous integration of the dcl1 knockout construct resulted in the replacement of the dcl1 coding gene (Pc21g06890) with a nourseothricin resistance cassette. Validation of homologous integration within the recipient strain Δdcl2 T1 was performed by Southern blotting and PCR of 20 μg BamHI digested genomic DNA with 32P-labeled complementary DNA probes of the 5'-flank (dark bar) of dcl1. Furthermore, PCR analyses, with a primer pair (indicated with grey arrows) surrounding the 3'-flank, which was used for homologous integration, validate the homologous
integration of the resistance cassette. In contrast to the tested genomic DNA of Δdcl1 T1 (+Control), the no-template controls (-Control) and PCRs with the recipient DNA (WT) show no PCR product. GeneRuler DNA Ladder (Thermo Scientific) was used as size standard. Homokaryotic transformants that show correct genomic integrations are marked with an asterisk.

**S1 Table.** Detailed list of predicted miRNA target sites.

**S2 Table.** Strains used in this study.

**S3 Table.** Cultivations used for RNA-Seq samples.

**S4 Table.** Components involved in small RNA biogenesis of *Penicillium chrysogenum* compared to five other fungi.

**S5 Table.** List of plasmids used in this study.

**S6 Table.** List of oligonucleotides used in this study.

**Acknowledgments**

We thank PD Dr. Minou Nowrousian (Lehrstuhl für Allgemeine und Molekulare Botanik, Ruhr-University Bochum, Germany) for her substantial advice in bioinformatics. We thank Drs. Ivo Zadra, Thomas Specht, and Hubert Kürnsteiner (Sandoz GmbH, Kundl, Austria) for their ongoing interest and support.

**Author Contributions**

Conceived and designed the experiments: TAD UK. Performed the experiments: TAD. Analyzed the data: TAD. Contributed reagents/materials/analysis tools: UK. Wrote the paper: TAD UK.

**References**

1. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014; 15:509–24. doi:10.1038/nrm3838 PMID: 25027649
2. Axtell MJ. Classification and comparison of small RNAs from plants. Annu Rev Plant Biol. 2013; 64:137–59. doi:10.1146/annurev-arplant-050312-120043 PMID: 23330790
3. Avesson L, Reimegård J, Wagner EGH, Söderbom F. MicroRNAs in Amoebozoa: Deep sequencing of the small RNA population in the social amoeba *Dictyostelium discoideum* reveals developmentally regulated microRNAs. RNA. 2012; 18:1771–82. doi:10.1261/rna.033175.112 PMID: 22875808
4. Kim V. Small RNAs: classification, biogenesis, and function. Mol cells. 2005; 19:1–15. PMID: 15750334
5. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet. 2009; 10:94–108. doi:10.1038/nrg2504 PMID: 19148191
6. Ghildiyal M, Seitz H, Horwich MD, Li C, Du T, Lee S, et al. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. Science. 2008; 320:1077–81. doi:10.1126/science.1157396 PMID: 18409677
7. Bologna NG, Voinnet O. The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. Annu Rev Plant Biol. 2014; 65:473–503. doi: 10.1146/annurev-arplant-050213-035728 PMID: 24579988

8. Chang S-S, Zhang Z, Liu Y. RNA interference pathways in fungi: mechanisms and functions. Annu Rev Microbiol. 2012; 66:305–23. doi: 10.1146/annurev-micro-092611-150138 PMID: 22746336

9. Nicolás FE, Ruiz-Vázquez RM. Functional diversity of RNAi-associated sRNAs in fungi. Int J Mol Sci. 2013; 14:15348–60. doi: 10.3390/ijms140815348 PMID: 23887655

10. Cogoni V, Macino G. Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature. 1999; 399:166–9. PMID:10335848

11. Fulci V, Macino G. Quelling: post-transcriptional gene silencing guided by small RNAs in Neurospora crassa. Curr Opin Microbiol. 2007; 10:199–203. PMID:17395524

12. Lee H-C, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, et al. Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. Mol Cell. 2010; 38:803–14. doi: 10.1016/j.molcel.2010.04.005 PMID: 20417140

13. Jiang N, Yang Y, Janbon G, Pan J, Zhu X. Identification and functional demonstration of miRNAs in the fungus Cryptococcus neoformans. PLoS One. 2012; 7:e52734. doi: 10.1371/journal.pone.0052734 PMID: 23300755

14. Zhou J, Fu Y, Xie J, Li B, Jiang D, Li G, et al. Identification of microRNA-like RNAs in a plant pathogenic fungus Sclerotinia sclerotiorum by high-throughput sequencing. Mol Genet Genomics. 2013; 287:276–82. doi: 10.1007/s00438-012-0678-8 PMID: 22988464

15. Chen R, Jiang N, Jiang Q, Sun X, Wang Y, Zhang H, et al. Exploring microRNA-like small RNAs in the filamentous fungus Fusarium oxysporum. PLoS One. 2014; 9:e104956. doi: 10.1371/journal.pone.0104956 PMID: 25141304

16. Lau SKP, Chow W-N, Wong AYP, Yeung JMY, Bao J, Zhang N, et al. Identification of microRNA-like RNAs in mycelial and yeast phases of the thermal dimorphic fungus Penicillium marneffei. PLoS Negl Trop Dis. 2013; 7:e2398. doi: 10.1371/journal.pntd.0002398 PMID: 23991243

17. Axtell MJ, Westholm JO, Lai EC. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol. 2011; 12:221. doi: 10.1186/gb-2011-12-4-221 PMID: 21564756

18. Wilczynska a, Bushell M. The complexity of miRNA-mediated repression. Cell Death Differ. 2015; 22:22–33. doi: 10.1038/cdd.2014.112 PMID: 25190144

19. Li L, Chang S, Liu Y. RNA interference pathways in filamentous fungi. Cell Mol Life Sci. 2010; 67:3849–63. doi: 10.1007/s00018-010-0471-y PMID: 20680389

20. Nakayashiki H. RNA silencing in fungi: Mechanisms and applications. FEBS Letters. 2005. p. 5950–7.

21. Drinnenberg I a, Fink GR, Bartel DP. Compatibility with killer explains the rise of RNAi-deficient fungi. Science. 2011; 333:1592. doi: 10.1126/science.1209575 PMID: 21921191

22. Specht T, Dahmann TA, Zadra I, Kümstein M. Evidence for Dicer-dependent RNA interference in the industrial penicillin producer Penicillium chrysogenum. Microbiology. 2009; 155:3946–56. doi: 10.1099/mic.0.032763-0 PMID: 19797363

23. Hoff B, Kück U. Evidence for Dicer-dependent RNA interference in the industrial penicillin producer Penicillium chrysogenum. Genome Announc. 2014; 2:e00577-14. doi: 10.1128/genomeA.00577-14 PMID: 25059858

24. Hoff B, Kamerewerd J, Sigl C, Zadra I, Kück U. Homologous recombination in the antibiotic producer Penicillium chrysogenum: strain DeltaPcku70 shows up-regulation of genes from the HOG pathway. Appl Microbiol Biotechnol. 2010; 85:1081–94. doi: 10.1007/s00253-009-2166-4 PMID: 19690852

25. Drinnenberg I, Fink GR, Bartel DP. Compatibility with killer explains the rise of RNAi-deficient fungi. Science. 2011; 333:1592. doi: 10.1126/science.1209575 PMID: 21921191

26. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into Arabidopsis Argonaute complexes is directed by the 5′ terminal nucleotide. Cell. 2008; 133:116–27. doi: 10.1016/j.cell.2008.02.034 PMID: 18342361

27. Friedländer MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, et al. Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol. 2008; 26:407–15. doi: 10.1038/nbt1394 PMID: 18392026
29. Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. 2012; 40:37–52. doi: 10.1093/nar/gkr868 PMID: 21911355

30. Braumann I, van den Berg M, Kempken F. Transposons in biotechnologically relevant strains of Aspergillus niger and Penicillium chrysogenum. Fungal Genet Biol. 2007; 44:399–414. PMID: 17881255

31. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003; 5:R1. PMID: 14709173

32. Bonnet E, He Y, Billiau K, Van de Peer Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics. 2010; 26:1566–8. doi: 10.1093/bioinformatics/btq233 PMID: 20430753

33. Dai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res. 2011; 39:W155–9. doi: 10.1093/nar/gkr319 PMID: 21622958

34. Nunes CC, Gowda M, Sailsbery J, Xue M, Chen F, Brown DE, et al. Diverse and tissue-enriched small RNAs in the plant pathogenic fungus, Magnaporthe oryzae. BMC Genomics. 2011; 12:288. doi: 10.1186/1471-2164-12-288 PMID: 21635781

35. Nicolas FE, Moxon S, de Haro JP, Calo S, Grigoriev I V., Torres-Martínez S, et al. Endogenous short RNAs generated by Dicer 2 and RNA-dependent RNA polymerase 1 regulate mRNAs in the basal fungus Mucor circinelloides. Nucleic Acids Res. 2010; 38:5535–41. doi: 10.1093/nar/gkq301 PMID: 20427422

36. Nolan T, Braccini L, Azzalin G, De Toni A, Macino G, Cogoni C. The post-transcriptional gene silencing machinery functions independently of DNA repression to express a LINE1-like retrotransposon in Neurospora crassa. Nucleic Acids Res. 2005; 33:1564–73. PMID: 15767281

37. Murata T, Kadotani N, Yamaguchi M, Tosa Y, Mayama S, Nakayashiki H. siRNA-dependent and-independent post-transcriptional cosuppression of the LTR-retrotransposon MAGGY in the phytopathogenic fungus Magnaporthe oryzae. Nucleic Acids Res. 2007; 35:5987–94. PMID: 17726046

38. Segers GC, Zhang X, Deng F, Sun Q, Nuss DL. Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. Proc Natl Acad Sci U S A. 2007; 104:12902–6. PMID: 17646660

39. Lein J. The Panlabs Penicillium strain improvement program. In: Vanek Z, Hostálek Z, editors. Overproduction of microbial metabolites. Butterworths; 1986. p. 105–40.

40. Nunes CC, Sailsbery JK, Dean R a. Characterization and application of small RNAs and RNA silencing mechanisms in fungi. Fungal Biol Rev. 2011; 25:172–80.

41. Felice K, Salzman D, Shubert-Coleman J, Jensen KP, Furneaux HM. The 5 terminal uracil of let-7a is critical for the recruitment of mRNA to Argonaute2. Biochem J. 2009; 422:329–41. doi: 10.1042/BJ20090534 PMID: 19508234

42. Catalanotto C, Pallotta M. Redundancy of the two dicer genes in transgene-induced posttranscriptional gene silencing in Neurospora crassa. Mol Cell Biol. 2004; 24:2536–45. PMID: 14993290

43. Ameres SL, Zamore PD. Diversifying microRNA sequence and function. Nat Rev Mol Cell Biol. 2013; 14:475–88. doi: 10.1038/nrm3611 PMID: 23860994

44. Böhm J, Hoff B, O’Gorman CM, Wolfers S, Klix V, Binger D, et al. Sexual reproduction and mating-type-mediated strain development in the penicillin-producing fungus Penicillium chrysogenum. Proc Natl Acad Sci U S A. 2013; 110:1476–81. doi: 10.1073/pnas.1217943110 PMID: 23307807

45. Wolfers S, Kamerewerd J, Sigl C, Zadra I, Künkler H, et al. Microarray hybridization analysis of light-dependent gene expression in the filamentous fungus Penicillium chrysogenum. Fungal Genet Biol. 2007; 44:1399–408. doi: 10.1016/j.fgb.2007.03.004 PMID: 17643775

46. Van den Berg MA, Albanger R, Albermann K, Badger JH, Daran J-M, Driessen AJM, et al. Genome sequencing and analysis of the filamentous fungus Penicillium chrysogenum. Nat Biotechnol. 2008; 26:1161–8. doi: 10.1038/nbt.1498 PMID: 18820685

47. Altatschul S, Gish W, Miller W. Basic local alignment search tool. J Mol Biol. 1990; 215:403–10. PMID: 2231712

48. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30:3059–66. PMID: 12136088

49. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 2011; 39:D225–10. doi: 10.1093/nar/gkq1189 PMID: 21109532

50. Kopke K, Hoff B, Kück U. Application of the Saccharomyces cerevisiae FLP/FRT recombination system in filamentous fungi for marker recycling and construction of knockout strains devoid of heterologous genes. Appl Environ Microbiol. 2010; 76:4664–74. doi: 10.1128/AEM.00670-10 PMID: 20472720
51. Windhofer F, Hauck K, Catcheside DEA, Kück U, Kempken F. Ds-like restless deletion derivatives occur in Tolypocladium inflatum and two foreign hosts, Neurospora crassa and Penicillium chrysogenum. Fungal Genet Biol. 2002; 35:171–82. PMID: 11848679

52. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 1987; 162:156–9. PMID: 2440339

53. Hoff B, Kamerewerd J, Sigl C, Mitterbauer R, Zadra I, Kurnsteiner H, et al. Two components of a velvet-like complex control hyphal morphogenesis, conidiophore development, and penicillin biosynthesis in Penicillium chrysogenum. Eukaryot Cell. 2010; 9:1236–50. doi: 10.1128/EC.00077-10 PMID: 20543063

54. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011; 17:10–2.

55. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. doi: 10.1186/gb-2009-10-3-r25 PMID: 19261174

56. Molnár A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature. 2007; 447:1126–9. PMID: 17538623

57. Lorenz R, Bernhart SH, Höner Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol. 2011; 6:26. doi: 10.1186/1748-7188-6-26 PMID: 22115189

58. Church G, Gilbert W. Genomic sequencing. Proc Natl Acad Sci. 1984; 81:1991–5. PMID: 6326095

59. Kim SW, Li Z, Moore PS, Monaghan a P, Chang Y, Nichols M, et al. A sensitive non-radioactive northern blot method to detect small RNAs. Nucleic Acids Res. 2010; 38:e98. doi: 10.1093/nar/gkp1235 PMID: 20081203

60. Gesing S, Schindler D, Nowrousian M. Suppression subtractive hybridization and comparative expression analysis to identify developmentally regulated genes in filamentous fungi. J Basic Microbiol. 2013; 53:742–51. doi: 10.1002/jobm.201200223 PMID: 22961396

61. Nowrousian M, Ringelberg C, Dunlap JC, Loros JJ, Kück U. Cross-species microarray hybridization to identify developmentally regulated genes in the filamentous fungus Sordaria macrospora. Mol Genet Genomics. 2005; 273:137–49. PMID: 15778868