1498–1512 *Nucleic Acids Research, 2015, Vol. 43, No. 3* *Published online 20 January 2015*
doi: 10.1093/nar/gku1387

ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome occupancy at human Transcription Start Sites

Claudia Carissimi[1,†], Ilaria Laudadio[1,†], Emanuela Cipolletta[1], Silvia Gioiosa[1], Marija Mihailovich[2], Tiziana Bonaldi[2], Giuseppe Macino[1,*] and Valerio Fulci[1]

[1]Dipartimento di Biotecnologie Cellulari ed Ematologia, 'Sapienza' Università di Roma, Rome 00161, Italy and [2]Department of Experimental Oncology, European Institute of Oncology (IEO), Milan 20139, Italy

Received September 19, 2014; Revised November 24, 2014; Accepted December 22, 2014

ABSTRACT

Argonaute (AGO) proteins have a well-established role in post-transcriptional regulation of gene expression as key component of the RNA silencing pathways. Recent evidence involves AGO proteins in mammalian nuclear processes such as transcription and splicing, though the mechanistic aspects of AGO nuclear functions remain largely elusive. Here, by SILAC-based interaction proteomics, we identify the chromatin-remodelling complex SWI/SNF as a novel AGO2 interactor in human cells. Moreover, we show that nuclear AGO2 is loaded with a novel class of Dicer-dependent short RNAs (sRNAs), that we called swiRNAs, which map nearby the Transcription Start Sites (TSSs) bound by SWI/SNF. The knockdown of AGO2 decreases nucleosome occupancy at the first nucleosome located downstream of TSSs in a swiRNA-dependent manner. Our findings indicate that in human cells AGO2 binds SWI/SNF and a novel class of sRNAs to establish nucleosome occupancy on target TSSs.

INTRODUCTION

The first RNA interference (RNAi) phenomena were reported ∼20 years ago in plants (1), fungi (2) and metazoa (3). Key players in this process are sRNAs (19–24 nt), the type III ribonuclease Dicer which is involved in sRNA processing, and a widely conserved family of proteins called Argonaute (AGO) (4). AGO proteins are primarily known for their cytoplasmic function in post-transcriptional gene silencing. In the cytosol, AGO proteins bind to a sRNA molecule and to other proteins (TNRC6, TRBP), giving rise to the RNA-Induced Silencing Complex (RISC) . The sRNA guides the recruitment of the RISC complex onto target RNA molecules by base-pair complementarity.

Growing evidence, however, indicates that AGO proteins can also regulate nuclear processes in association with sRNAs (5). These mechanisms have been well-characterized in *Schizosaccharomyces pombe* where AGO proteins participate to the assembly of heterochromatin at centromeric regions via histone methylation (6), and in plants where RNAi directs establishment, spread and removal of DNA methylation (7). In the past few years, several studies have reported that AGO proteins exert pivotal functions also in the nuclei of mammalian cells. In particular, AGO1 and/or AGO2 in association with exogenous and/or endogenous sRNAs complementary to genomic target regions repress (8,9) or activate gene expression (10,11), control alternative splicing (12,13) and DNA repair (14,15).

The SWI/SNF (switch/sucrose non-fermentable) complex is a highly conserved multi-subunit ATP-dependent chromatin remodelling complex which alters the structure and positioning of nucleosomes, thereby modulating the access of regulatory proteins to DNA. In mammalian cells, the canonical SWI/SNF complex contains one of the two mutually exclusive ATPases, BRM (SMARCA2) or BRG1 (SMARCA4), a 'core' set of subunits consisting of BAF155 (BRG1-associated factor or SMARCC1), BAF47 (hSNF5 or INI1) and BAF170 (SMARCC2), as well as four to eight other accessory subunits (16). Through chromatin remodelling, the SWI/SNF complex controls transcription, DNA repair, recombination and chromosome segregation thus affecting a variety of biological processes such as cell differentiation, proliferation, development and malignant transformation (17).

Recently, genome-wide ChIP-Seq analyses conducted on different murine and human cell lineages have shown that a considerable fraction of SWI/SNF binding sites resides in functional genomic regions such as TSSs, enhancers, CCCTC-binding factor (CTCF)-bound loci and many re-

*To whom correspondence should be addressed. Tel: +39 6 4457731; Fax: +39 6 4457731; Email: claudia.carissimi@uniroma1.it
Correspondence may also be addressed to Giuseppe Macino. Tel: +39 6 4457731; Fax: +39 6 4457731; Email: giuseppe.macino@uniroma1.it
†These authors contributed equally to the paper as first authors.

gions occupied by RNA Polymerase II (18–20). Consistently, the SWI/SNF complex is capable of facilitating both gene activation and repression probably by establishing and maintaining the nucleosome landscape around TSSs (21).

Here, by using a SILAC-based interactomics approach we identify the SWI/SNF complex as a novel interactor of AGO2 in human cells. Moreover, we provide evidence of the existence of a novel class of AGO2-associated sRNAs, that we termed 'swiRNAs', which map nearby TSSs bound by SWI/SNF and are processed in a Dicer-dependent manner. We further demonstrate that AGO2 depletion results in altered occupancy of nucleosome +1 downstream of these TSSs. Our data provide the first evidence of a direct crosstalk between AGO2 and the chromatin remodelling machinery in human cells, suggesting a novel mechanism of epigenetic regulation.

MATERIALS AND METHODS

Cell culture and transfection

HeLa S3, Jurkat and HEK293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) foetal bovine serum, 2 mM L-glutamine and penicillin–streptomycin. HCT116 WT and Dicer^{Ex5} cells (22) were grown in McCoy's 5A medium supplemented with 10% (v/v) foetal bovine serum, 2 mM L-glutamine and penicillin–streptomycin.

Transfections were done with 10 nM siRNAs (siAGO2: a pool containing the following siRNAs GCAGGACAAAGAUGUAUAA[dT][dT] (23) and CGUCCGUGAAUUUGGAAUCAU[dT][dT] (Sigma); siCTRL: AGCUUCAUAAGGCGCAUGC[dT][dT]) for 4 days using INTERFERin® as transfecting agent according to the manufacturer's instructions (Polyplus Transfection).

RNA isolation and RT-qPCR

RNA from HeLa S3, HCT116 WT and Dicer^{Ex5} cells was isolated using TriReagent according to the manufacturer's protocol (Sigma).

For RT-qPCR, RNA was retrotranscribed using Enhanced Avian Reverse Transcriptase (Sigma), and specific reverse primers. Quantitative real-time PCR was performed using SensiMix SYBR & Fluorescein Kit (Bioline) with Biorad iCycler. Quantification was normalized to the small nucleolar RNA U44.

Primer sequences are available under request.

Western blot

For western blot analyses the following antibodies were used: anti-AGO1 (4B8, Ascenion), anti-AGO2 (11A9, Ascenion), anti-GAPDH (14C10, Cell Signaling technology), anti-histone H1 + core proteins (F152.C25.WJJ, Millipore) anti-β-tubulinI (SAP.4G5, Sigma), anti-BRG1 (H-88, Santa Cruz), anti-BAF155 (Abcam), anti-INI1 (Bethyl Laboratoties), goat-anti mouse and anti-rabbit IgG–HRP conjugated (Bio-Rad), anti-rat IgG–HRP conjugated (Jackson).

Chromatin binding assay

Chromatin binding assay was performed as previously described (24). The equivalent of 1×10^6 cells of all fractions were analysed by western blotting.

Preparation of total, nuclear and cytosolic cell extract

For total cell extract, cells lysed in immunoprecipitation (IP)-buffer (150 mM KCl; 25 mM NaCl; 2 mM EDTA; 0.5% NP40; 0.5 mM DTT; protease inhibitor (Sigma)) for 20 min on ice. Lysate was clarified by centrifugation at 16 000 g for 10 min at 4°C.

Cytosolic and nuclear extracts were prepared as following: cells were resuspended, firstly, in one volume of ice-cold buffer I (0.3 M sucrose in 60 mM KCl; 15 mM NaCl; 5 mM $MgCl_2$; 0.1 mM Ethylene Glycol Tetraacetic Acid; 15 mM Tris–HCl pH 7.5; 0.5 mM DTT; protease inhibitors), then added one volume of ice-cold buffer II (0.3 M sucrose in 60 mM KCl; 15 mM NaCl; 5 mM $MgCl_2$; 0.1 mM EGTA; 15 mM Tris–HCl pH 7.5; 0.5 mM DTT, 0.4% NP-40, protease inhibitors), and placed on ice for 10 min. Afterwards, cell lysate was layered on 24 ml of a sucrose cushion (1.2 M sucrose in 60 mM KCl; 15 mM NaCl; 5 mM $MgCl_2$; 0.1 mM EGTA; 15 mM Tris–HCl pH 7.5; 0.5 mM DTT) and centrifuged in a pre-chilled swing-out rotor at 10 000 g for 20 min at 4°C. The upper phase containing the cytoplasmic fraction was collect and the pellet containing cell nuclei was lysed in IP-buffer and clarified.

Immunoprecipitation and co-immunoprecipitation

Antibodies anti-AGO1 (4B8, Ascenion); anti-AGO2 (11A9, Ascenion); anti-BAF155 (Abcam)) and isotype-matched IgG (as mock IP) were coupled to protein-G-sepharose beads (Sigma).

Whole-cell lysate (40-80 $\times$ 10^6 cells) or nuclear fraction (60-160 $\times$ 10^6 cells) were incubated overnight at 4°C with antibodies-coupled beads. IP were washed once with IP-buffer and three times with IP-wash buffer (50 mM Tris–HCl, pH 7.5; 300 mM NaCl; 5 mM $MgCl_2$; and 0.05% nonidet P-40). Proteins were eluted from beads by boiling for 5 min in SDS-PAGE sample buffer (Sigma).

DNase I and RNase A treatment

HeLa S3 whole-cell lysate was incubated with or without DNase I (Roche) or RNaseA (Sigma) for 1 h at 4°C. EDTA was then added to a final concentration of 10 mM to stop reaction. Lysates treated with or without DNase I or RNaseA were immunoprecipitated using anti-AGO2 antibody or IgG, following the protocol previously described.

Nuclear RNA-immunoprecipitation (RIP)

Nuclei of HeLa S3, Jurkat, HCT116 WT and Dicer^{Ex5} cell lines were lysed in IP-buffer supplemented with RNasin (Promega). Nuclear lysates were pre-cleared in the presence of protein G-sepharose beads for 2 h at 4°C.

Antibodies and isotype-matched IgG (mock IP) were coupled to protein G-sepharose beads in IP-buffer containing 1 mg/ml heparin (Sigma). The pre-cleared lysate

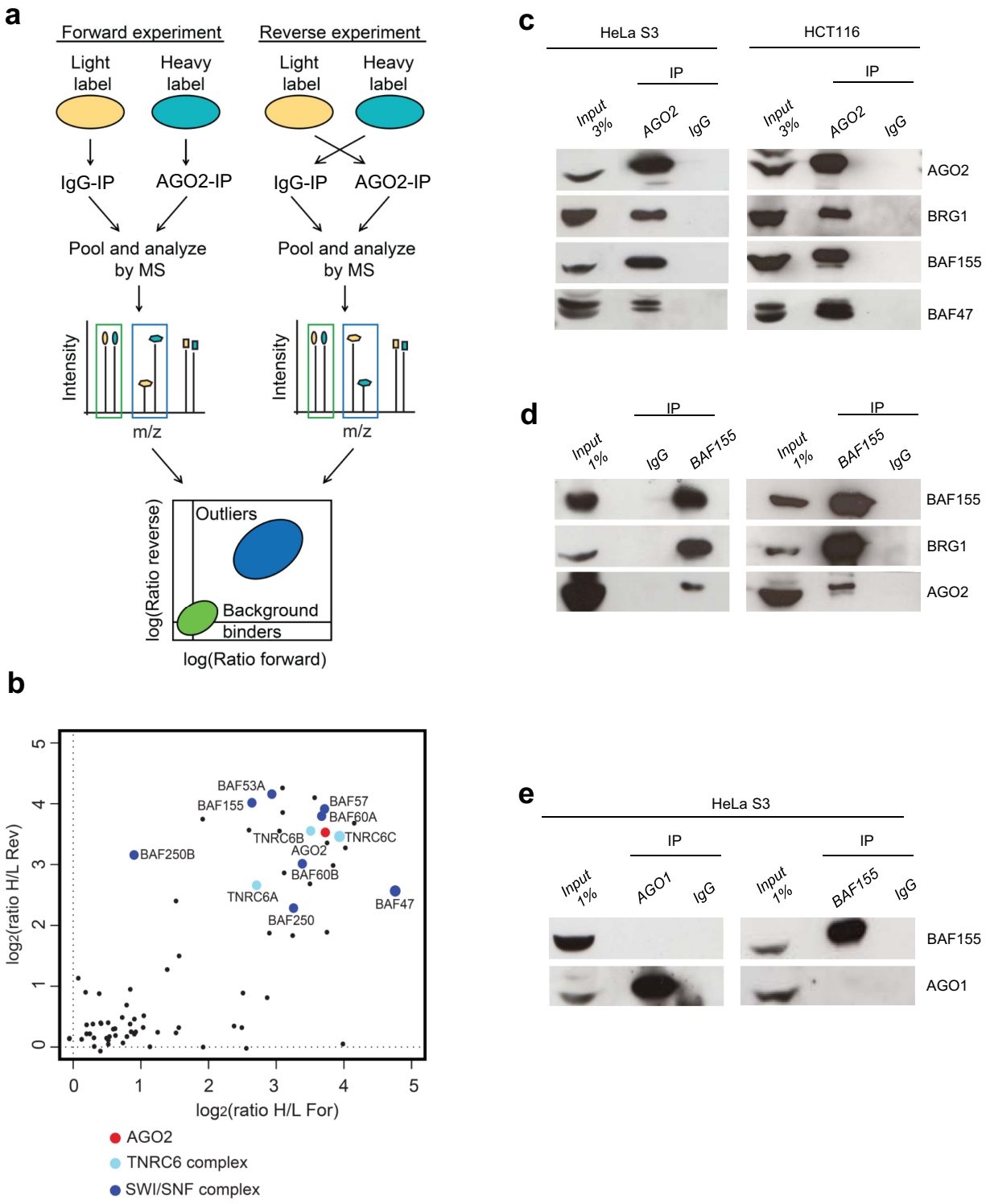

Figure 1. AGO2 interacts with components of SWI/SNF complex in human cell lines. (**a**) Schematic view of SILAC-based analysis of the AGO2 interactome, carried out in forward (for) and reverse (rev) setup (upper panel). The scatter plot based on protein SILAC-ratios (Heavy/Light) in forward and reverse experiments (x and y axes, respectively), shows that the outliers (true binders; blue) are in the top right quadrant, while background proteins (unspecific binders; green) cluster around protein ratio 1:1 (0 on a logarithmic scale). (**b**) Experimental scatterplot of H/L protein ratios obtained in the forward and reverse AGO2-IP SILAC experiment with HeLaS3 lysates. Components of SWI/SNF and TNRC6 complex are highlighted in dark and light blue respectively, while the bait is in red. (**c**) HeLaS3 and HCT116 cell extracts were immunoprecipitated (IP) with AGO2 antibody or isotype-matched immunoglobulins (IgG) as a negative control. (**d**) HeLaS3 and HCT116 cell extracts were immunoprecipitated with BAF155 antibody or IgG as a negative control. (**e**) AGO1 does not interact with SWI/SNF complex in human cell lines. HeLaS3 cell extract was immunoprecipitated with AGO1 or BAF155 antibody (with IgG as a negative control). Input and IPs were immunoblotted with the indicated antibodies.

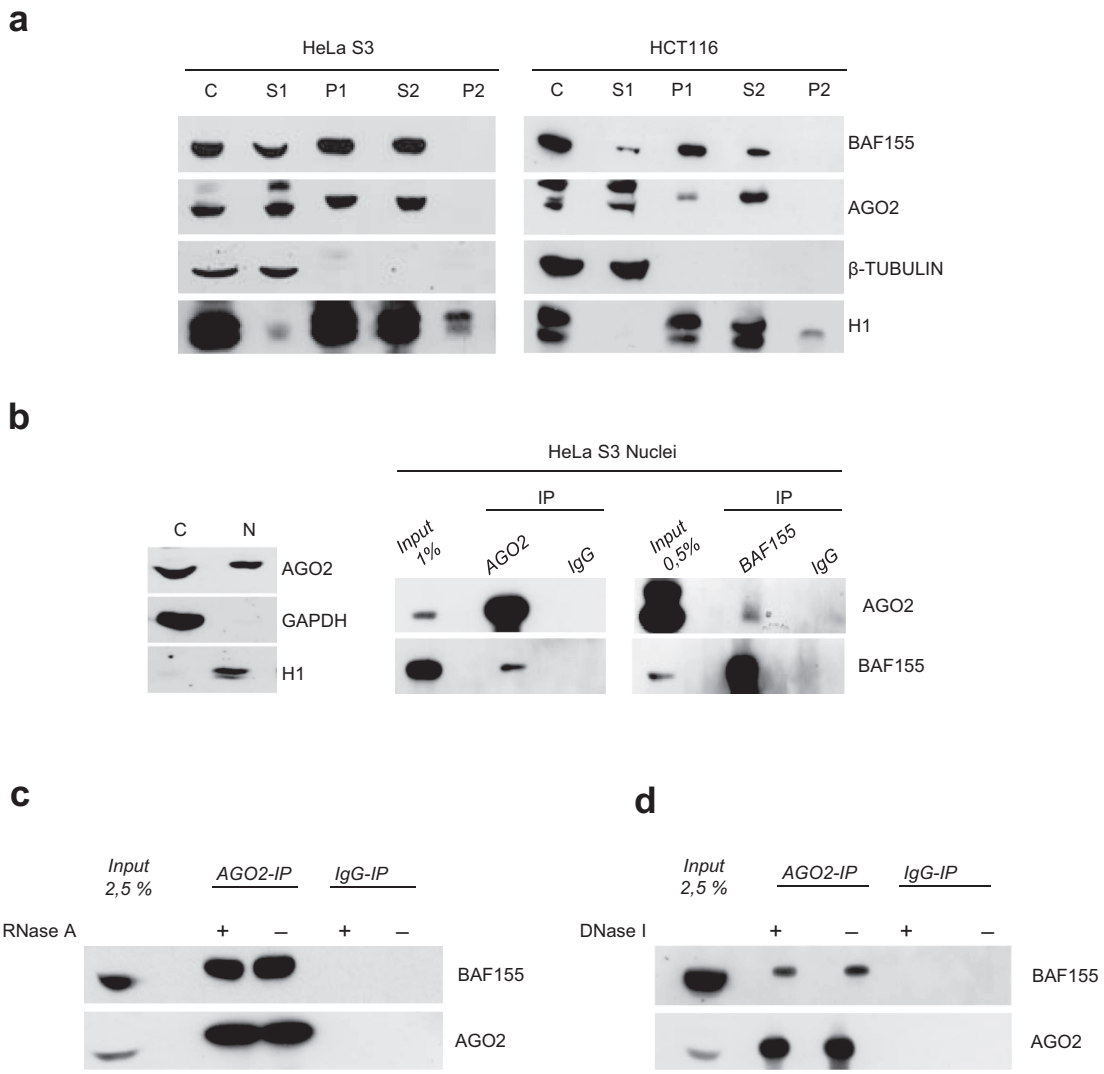

Figure 2. AGO2 and SWI/SNF complex are associated with chromatin and interact in the nucleus in a RNA- and DNA-independent manner. (**a**) Whole cell extract (C), triton-soluble protein (S1), chromatin-bound, nuclear-matrix-bound and insoluble protein (P1), DNase-released chromatin associated protein (S2) and insoluble, cytoskeletal and nuclear matrix protein (P2) fractions corresponding to the same amount of HeLaS3 or HCT116 cells were analysed by western blot for the presence of indicated proteins. β-tubulin serves as a chromatin-unbound marker whereas histone H1 as chromatin-bound marker. (**b**) Cytosolic and nuclear fractions of HeLaS3 cells were prepared. Fractionation efficiency was verified by western blot for the presence of GAPDH (cytosolic marker) and H1 (nuclear marker) (Left panel). Nuclear lysates were immunoprecipitated with indicated antibodies or IgG as a negative control (right panel). (**c**) Total cell extract pre-treated or not with RNase A (left panel) or DNase I (right panel) were immunoprecipitated with AGO2 antibody or IgG as a negative control. Input and IPs were immunoblotted with the indicated antibodies.

(150×10^6 nuclei for AGO2–RIP or 350×10^6 cells for AGO1–RIP) were incubated with antibody-coupled beads overnight at 4°C. RNA from Input and IP samples was extracted by phenol:chloroform:isopropyl alcohol, after DNaseI- and proteinase K-treatment, and precipitated in ethanol.

Cell culture and SILAC (stable isotope labelling by amino acids in cell culture) labelling

HeLa S3 and Jurkat cell lines were grown in SILAC DMEM (PAA) and SILAC RPMI (Sigma), respectively. The SILAC media were supplemented with 10% dialysed foetal bovine serum (PAA), non-essential amino acids L-glutamine and L-leucine (Sigma), penicillin/streptomycin and either unlabelled amino acids ('Light', L-lysine and L-arginine, Sigma) or isotope-labelled amino acids ('Heavy', L-lysine-$^{13}C_6$–$^{15}N_2$ and L-arginine–$^{13}C_6$–$^{15}N_4$, Sigma/Isotec). Cells were adapted for at least six generations to achieve complete incorporation of the isotope-labelled amino acids before they were propagated to a scale needed for IP (350×10^6 cells for HeLa S3 cells and 150×10^6 cells for Jurkat cells). AGO2_IP was firstly performed using the lysate from 'Heavy' cells as input, whereas mock IP (IgG_IP) was performed using the the 'Light' cellular lysate. In a second experiment, a labelling swap was performed. IP was performed as previously described with the following differences: equal amounts of proteins from Heavy and Light lysates were used for IP. Heavy and Light IP were mixed 1:1 before elution. The elution was performed twice by compe-

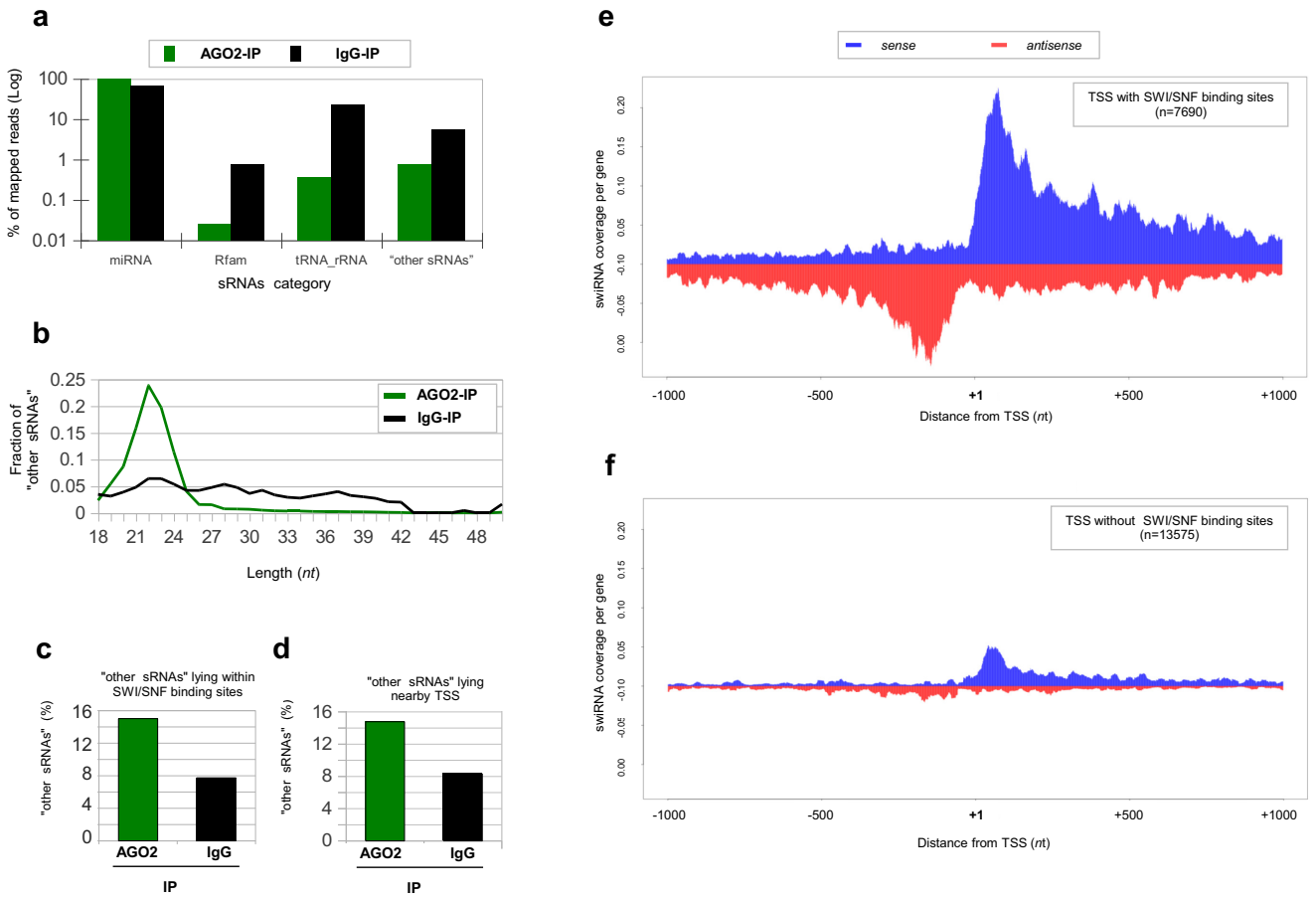

Figure 3. Nuclear AGO2 is associated with a novel class of sRNAs arising from TSSs bound by SWI/SNF complex (swiRNAs). (**a**) Genomic annotation of sRNA-Seq data (hg19). The histograms illustrate the sRNA classes identified in AGO2- or IgG-IP samples. The main RNA classes in the AGO2-IP sample are represented by miRNAs followed by 'other sRNAs'. (**b**) Size distribution of AGO2- and IgG-IP 'other sRNAs'. The fraction of 'other sRNAs' in AGO2- or IgG-IP samples and corresponding length (*nt*) are plotted. (**c** and **d**) Percentage of 'other sRNAs' in AGO2- or IgG-IP samples mapping within SWI/SNF biding sites (ENCODE ChIP-seq data) (**c**) or within 1 kb of TSSs of expressed genes (TSSs were defined by mRNA-seq) (**d**). (**e** and **f**) Average per gene coverage of AGO2-associated 'other sRNAs' around TSSs with a SWI/SNF binding site (ENCODE ChIP-seq data) within ±1 kb (**e**) and around TSSs without SWI/SNF binding sites (**f**). Average per gene coverage is defined as the sum of sRNAs mapping to a given position (relative to TSS) divided by number of TSSs analysed. Only 'other sRNAs' lying within clusters of less than 50 molecules were considered. Red and blue represent sRNAs in the sense and anti-sense orientation with respect to gene transcription, respectively.

tition using a molar excess of a synthetic AGO2 peptide (3 μg/10 × 10^6 cells; GeneScript) in elution buffer (300 mM NaCl, 50 mM Tris–HCl pH 8, 0.01% NP-40, 5 mM $MgCl_2$) for 1 h at 22°C, 750 rpm. The eluates were cleaned up with Micro Bio-Spin Chromatography Columns (Bio-Rad).

SDS-PAGE and protein digestion

Proteins eluted from the beads in the IP experiments were separated on a 4–12% gradient Bis–Tris mini gel (Invitrogen). After Coomassie staining (Colloidal Blue Staining Kit, Invitrogen), each lane was cut in seven slices and trypsin digested according to a previously described protocol (25). Briefly, upon destaining by incubation with 50% acetonitrile/25 mM ammonium bicarbonate (NH_4HCO_3) solution, followed by 100% acetonitrile for dehydration, gel pieces were incubated with 10 mM dithiothreitol in 50 mM NH_4HCO_3 for 60 min at 56°C for cysteine reduction followed by treatment with 55 mM iodoacetamide in 50 mM NH_4HCO_3 for 45 min at room temperature, in dark, for cysteines alkylation. After several rounds of washing with 50 mM NH_4HCO_3 and dehydrating with 100% acetonitrile, trypsin was added to the gel slices and incubated overnight at 37°C. The next day tryptic mixture was acidified with 2 μl of 50% trifluoroacetic acid solution and collected. Peptides were further eluted by 30% acetonitrile, 3% trifluoroacetic acid and with 100% acetonitrile. Eluted peptides were concentrated in a speed-vacuum centrifuge and solubilized in 100 μl of 0.1% formic acid, desalted and concentrated using reverse-phase C18 handmade nano-columns (StageTips) (26). Samples were eluted from StageTips with 80% acetonitrile, lyophilized and re-suspended in 7 μl of 0.1% formic acid for LC–MS/MS analysis.

Liquid chromatography and tandem mass spectrometry (LC-MS/MS)

Nano-columns (15 cm long with 75 μm inner diameter and 350 μm outer diameter) were prepared by packing C18 resin (ReproSil C18-AQ, 3 μm Dr Maisch, Germany). The 5 μl

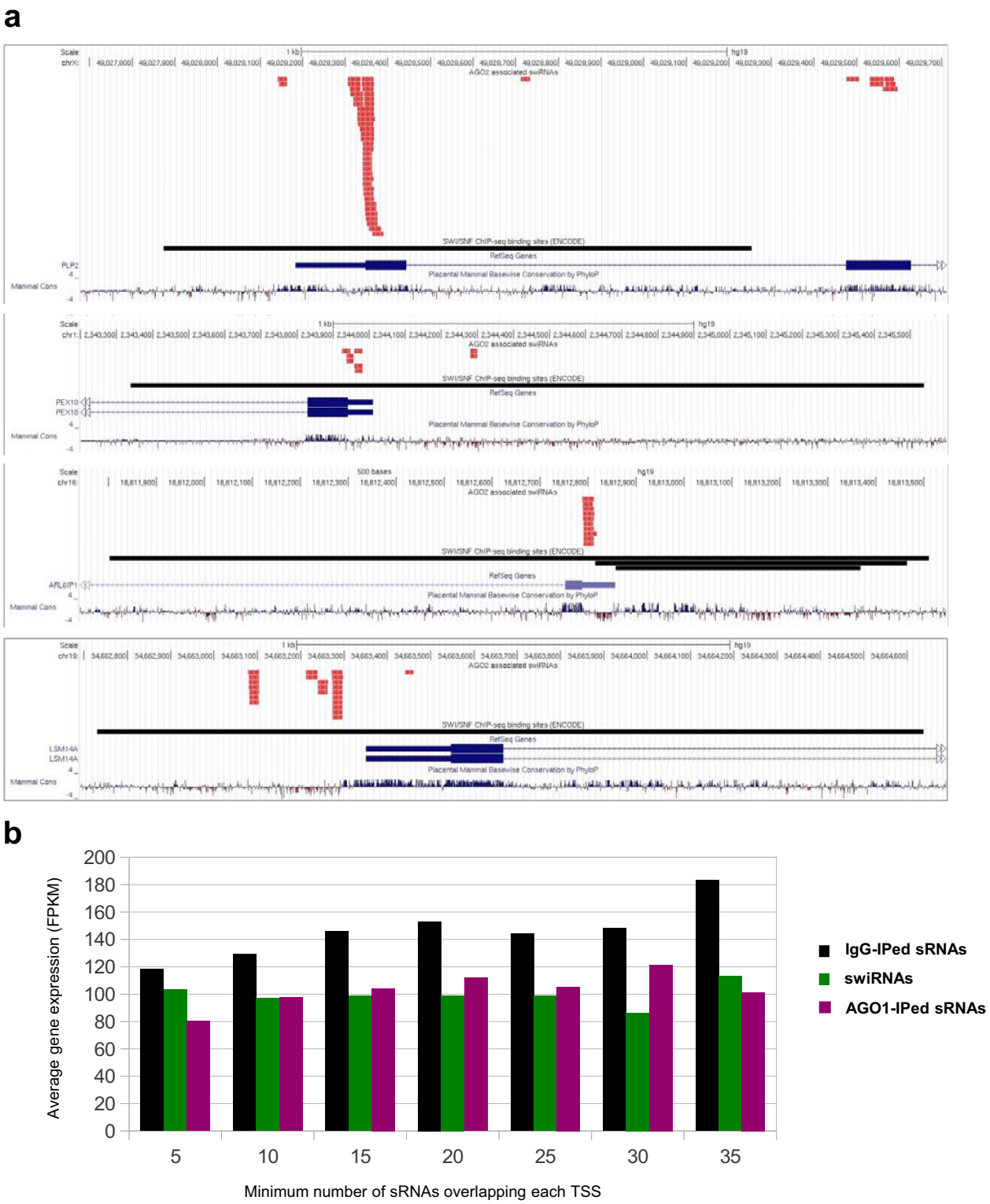

Figure 4. The number of AGO2-associated swiRNAs overlapping each TSS does not correlates with gene expression level. (**a**) Genome browser (USCS) view of representative clusters of AGO2-associated swiRNAs in HelaS3 cells. swiRNAs map to TSSs overlapped by SWI/SNF binding sites. Reads corresponding to AGO2-associated swiRNAs are red-coloured and orientation respect to reference genome is indicated for each read. SWI/SNF binding sites are black-coloured (BAF155, BAF170, BRG1, BAF47. Data from ENCODE). (**b**) Bar height represents average gene expression of loci with TSS (±150 nt) overlapped by at least the indicated number of AGO2-bound swiRNAs (green), IgG-IPed 'other sRNAs' (black) and AGO1-associated 'other sRNAs' (purple) in HeLaS3 cells.

of desalted peptides were injected in the nano-column at 500 nl/min and separated by nanoflow liquid chromatography on an Agilent 1100 Series LC system (Agilent Technologies), on line with a LTQ-FT Ultra mass spectrometer (Thermo Scientific). The gradient used for peptide separation was performed at a flow rate of 250 nl/min and was 0–36% over 120 min, followed by gradients of 36–60% for 10 min and 60–80% for 5 min. Solvent A was 0.1% formic acid and 5% acetonitrile in ddH_2O, and solvent B was 95% acetonitrile with 0.1% formic acid. The mass spectrometer was controlled by Xcalibur software (Thermo Scientific) and operated in data-dependent acquisition mode, to automatically switch between MS and MS/MS scans. Survey full scan MS spectra were acquired by the FT detector from 200 to 1650 m/z, then the five most intense ions with charge 2+ to 4+ were selected for fragmentation in the linear ion trap by collision induced dissociation. Collision energy was set to 35 eV. In the LTQ-FT Ultra, full scan MS spectra were acquired at a target value (AGC) of 10^6 ions, with a resolution equal to 10^5 at 400 m/z.

Protein identification and quantification

Acquired spectra were matched in the human IPI database version 1301 by MaxQuant (27) version 1.3.0.05, which performs peak list generation, protein identification by Andromeda search engine (28) and protein quantitation based on SILAC. Cysteine carbamidomethylation and methionine oxidation were set as fixed and variable modification respectively and a maximum of two missed cleavages were allowed. Mass tolerance was set to 10 ppm for MS spectra and 0.5 Da for MS/MS spectra. Protein and peptide false discovery rate were set to 1%.

For each cell line, the forward and reverse experiments were analysed together using the 'match between runs' option to increase the number of identified peptides. Upon MaxQuant analysis, proteins identified by at least two peptides, of which at least one unique, were considered as high quality identification and further analysed if their ratio was calculated based on at least two counts (ratio count (RC) > 1). In the case of redundancy in protein identification (peptides matching to multiple protein groups), we report the leading protein in the 'protein groups' output table from MaxQuant.

Isolation of nucleosomal DNA by micrococcal nuclease (MNase) digestion

Digestion of chromatin from untreated, siCTRL- or siAGO2-treated HeLa S3 cells (2×10^6) was performed with 50 U of MNase (New England Biolabs) in 300 μl of permeabilization buffer (15 mM Tris–HCl pH 7.4, 300 mM sucrose, 60 mM KCl, 15 mM NaCl, 4 mM $CaCl_2$, 0.5 mM EGTA, 0.2% NP-40, 0.5 mM β-mercaptoethanol) for 20 min at 37°C. The reaction was stopped by addition of 300 ul of stop buffer (50 mM Tris–HCl pH 8, 20 mM EDTA, 1% SDS) for 2 min on ice. RNA was degraded with 75 μg RNase A for 1 h at 37°C and cellular proteins digested with 30 μg proteinase K for 1 h at 55°C. Nucleosomal DNA was purified by phenol:chloroform extraction and precipitated in ethanol.

Library construction and sequencing

Library construction and sequencing was performed by the Institute of Applied Genomics (IGA) Technology Services (Italy). Samples were sequenced on an Illumina HiSeq2000.

mRNA sequencing analysis

Total mRNA was purified from HeLaS3 cells. The data were submitted to ArrayExpress database and are available under the following accession number: E-MTAB-2128.

Reads were aligned using TopHat 2.0.9 (29) and transcripts were assembled with cufflinks (v 2.1.1) (30) with the following options:

-g iGenome_hg19_genes.gtf -M mask_hg19.gtf -b hg19.fa

where:

iGenome_hg19_genes.gtf contains annotation of known transcripts (iGenome UCSC hg19);

mask_hg19.gtf contains genomic coordinates of annotated rRNAs and tRNAs;

hg19.fa is the fasta file containing sequence of the human genome hg19.

Transcripts were merged using cuffcompare (2.1.1) (30) using the -R option, to exclude from final report any transcript which was not overlapped by any fragment in our RNA-seq dataset. We therefore ended up with a list of genes expressed in HeLaS3 cells under our growth conditions. All analyses conducted using TSS coordinates (sRNAs coverage, nucleosome occupancy) were restricted to the 21 265 loci expressed in HeLaS3 cells under our growth conditions. No further filtering based on the level of expression of mRNAs was applied. Expression values in FPKM summarized at gene level were used to compute average expression values reported in Figure 4b.

Small and long RNA sequencing analysis

The data were submitted to ArrayExpress database and are available under the following accession numbers:

HeLaS3 nuclear AGO2 associated sRNAs: E-MTAB-2084

HeLaS3 nuclear AGO1 associated sRNAs: E-MTAB-2129

HCT116 and HCT116 $Dicer^{Ex5}$ nuclear AGO2 associated sRNAs: E-MTAB-2056

Reads quality was checked using FastqC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)

Sequences were quality trimmed (H.Li: https://github.com/lh3/seqtk) and adapters were removed using cutadapt v 1.0 (31)

Sequences were iteratively aligned to different databases to identify RNAs which did not belong to any known AGO2-associated class of RNA. After each alignement round reads with no alignments were used in the next step. Alignements were performed using Bowtie 0.12.7 (32) with the following options: -n 0 -l 18 (requiring no mismatches in the first 18 nt of the reads).

In a first step, reads were aligned to human miRNA hairpin precursors miRbase (v20.0) (33) to remove all reads aligning to miRNA precursors.

In a second step, unaligned reads were aligned against rRNA and tRNA database to remove any reads aligning to tRNA and ribosomal RNA.

In a third step, unaligned reads were aligned against a manually cured Rfam database (34). Such manually cured Rfam database includes all sequences included in Rfam with the exception of those annotated as lincRNAs.

Finally, unaligned reads were mapped against the human genome (hg19) using bowtie 0.12.7 (32) with the following options: -n 0 -l 18 -m 1. RNAs mapping to a single locus on hg19 at this step were called 'other sRNAs' and further analysed.

'Other sRNAs' alignements were converted into BED format for both AGO2 sample and IgG control (AGO2_sRNAs.BED and IgG_sRNAs.BED) and further analysed using the bedtools package (35) as follows:

'Other sRNAs' were divided in clusters (requiring same strandness) using mergeBed program (mergeBed -s); most clusters consisted of less than 50 sRNAs, while a few hundred clusters consisted of tens of thounsands of identical reads. These high-copy clusters likely represent novel miRNAs or PCR artifacts. We therefore decided to focus our attention on sRNAs in low copy clusters (<50 sRNAs) and removed from AGO2_sRNAs.BED and IgG_sRNAs.BED reads lying into High-copy clusters (>50 reads).

To compute the overlap between sRNAs and other genomic features we calculated the fraction of 'other sRNAs' with at least one nt overlap with the feature of interest (e.g. CpG islands) using intersectBed tool. Exact binomial test was used to compute *P*-values, using the frequency of 'other sRNAs' overlapping a given feature as the expected frequency (null hypothesis).

coverageBed tool was used with the -s and -S options to compute coverage of sense 'other sRNAs' and antisense 'other sRNAs' around each TSS. The TSS coordinates were obtained by analysis of HeLaS3 mRNA-seq (see above). For analysis of HCT116 and HCT116 Dicer^{Ex5} cells TSS of human genes were retrieved from ENSEMBL human gene annotation (version 69) (36). Occupancy profiles around each TSS were piled up and normalized by the total number of 'other sRNAs' to obtain the cumulative profile depicted in Figures 3e, f and 5a, b.

In order to find putative dsRNA precursors of swiRNAs, HeLaS3 long nuclear RNA-seq data sets were downloaded from ENCODE (37) experiment matrix (GSM765403, GSM767848).

The long Poly-A plus (two replicates) and Poly-A minus (two replicates) datasets were pooled together to increase the sequencing depth. swiRNAs coordinates were compared to long RNA coordinates with intersectBed tool requiring that long RNA molecules derived from sense/antisense transcription completely embedded the mature swiRNA.

Nuclear AGO1-associated sRNAs were analysed using the same methodology outlined for AGO2.

4842 clusters of 'other sRNAs' consisting of >50 reads were identified (3.9% of clusters) and reads lying within these clusters were discarded from subsequent analyses.

To assess whether swiRNAs could arise from miRNA-like hairpin precursors we extended swiRNA sequences by 70 nt on both sides and predicted secondary structures of the resulting sequences with RNAfold. The obtained structures were parsed by a perl script requiring: (i) that at least 75% of the swiRNA nucleotides were paired in the resulting structure; (ii) that all paired nucleotides within the resulting structure were belonging to a single stem structure; (iii) that at least one of the three last nucleotides at each end of the swiRNA was paired in the resulting structure. Sequences satisfying all these conditions were scored as potential miRNA-like hairpins. Human mature miRNAs from miRbase v 20 were used as a positive control. Dinucleotide shuffling of swiRNAs and mature miRNA were used as negative controls.

ChIP sequencing analysis

To compute the coverage of BRG1 and BAF47 ChIP experiments raw data were downloaded from ENCODE (GEO accession GSM935635, GSM935511), aligned to hg19 human genome using bowtie v 0.12.7. Average size of ChIP fragments was calculated using macs14 software. Each mapped read (in .bed format) was extended to the average ChIP fragment length. Coverage at each nt around each TSS was computed using coverageBed (35).

MNase sequencing analysis

592534546 reads were sequenced for AGO2 knock-down sample

711051028 reads were sequenced for control sample

The data were submitted to ArrayExpress database and are available under the following accession numbers: E-MTAB-1844.

Sequence quality was assessed using FastQc v 0.10.1. (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)

Paired End reads were aligned using bowtie v 0.12.7 (32) with options -n 1 -l 30 -m 1.

Only properly paired reads were further analysed. MNase digestion is known to produce fragments of variable size; however, shorter fragments (<100 nt) are likely to represent the DNA footprint of transcription factors, polymerases and other DNA binding proteins. The DNA footprint of a nucleosome is expected to be about 146 nt long. However, it is known that at low efficiency MNase may occasionally nick the 146 nt long nucleosomal DNA at about 10 nt (one DNA helix turn) into the nucleosome at each end, giving rise to a small satellite population of 127 nt fragments which do represent nucleosomes. DNA fragments longer than 200 nt are unlikely to represent bona fide nucleosome footprints. We therefore selected paired end reads with a size between 100 and 200 nt.

After such selection, we obtained:

206561235 fragments for AGO2 knock-down sample

240239066 fragments for control sample

As a first control we checked size distribution of these reads, and as expected we observe two sharp peaks at 128 and 147 nt (Supplementary Figure S6). Importantly, the size distribution of control and AGO2 knock-down samples are extremely similar. This similarity confirms that the two DNA samples were digested in comparable conditions. Over-digestion of a sample compared to the other would result in a higher 128 nt peak and a lower 147 nt peak.

Nucleosome occupancy at the genomic loci of interest was computed using coverageBed and a custom perl script to determine cumulative occupancy over multiple loci. Occupancy was normalized by the total number of sequenced nucleotides in each library taking into account properly paired read pairs with a length between 100 and 200.

HeLaS3 TSS coordinates were based on RNA-seq analysis of mRNAs (see below). Occupancy profile at TSS overlapped by at least 30 swiRNAs (Figure 6c) was smoothed using supsmu function in R statistical environment.

Nucleosome +1 occupancy was defined as the sum of the coverage at each nucleotide position in the interval between nt +100 and nt +300 relative to TSS. For each group of genes that were selected based on the minimum number of swiRNAs mapping within ±150 nt of TSS, a paired Wilcoxon test was performed to compute *P*-value of the difference observed between average coverage at nucleosome +1 in AGO2 knock-down and CTRL siRNA cells. For each group of *n* genes (whose TSS were overlapped by at least m swiRNAs) 10 000 random permutations were performed to estimate FDR. In each permutation, *n* random genes were chosen (among the 21 265 genes expressed in HeLaS3 cells) and the *P*-value was computed comparing occupancy at nucleosome +1 with the same procedure outlined above for the real case.

GO term analysis

The Gorilla (38) online web tool was used to compute enriched GO terms in the ranked list of HeLaS3 expressed genes. The list was ranked in decreasing order according to the number of swiRNAs mapping within ±150 nt of TSS.

RESULTS

AGO2 interacts with SWI/SNF complex in the nucleus of human cells

In order to identify novel AGO2 interactors in human cells, we performed a SILAC-based proteomic analysis of proteins that co-immunoprecipitate with AGO2 in HeLaS3 and Jurkat cell lines. Two independent biological replicates were analysed for each cell line, with an isotope-labelling swap (Figure 1a). As expected, well-known AGO2 interactors (such as TNRC6 complex (39) and HSP90 (40)) were enriched in AGO2 immunoprecipitates (AGO2-IP) relative to mock IP (IgG-IP). Surprisingly, the complete set of SWI/SNF chromatin remodelling complex subunits was also significantly enriched in AGO2-IP as measured by their SILAC H/L protein ratio (Figure 1b, Supplementary Figure S1a–d and Table S1). Comparable results were obtained in both cell lines (Supplementary Table S2). We confirmed the enrichment of the core components of SWI/SNF complex (BAF155, BRG1, BAF47) in AGO2-IP by western blot in HeLaS3, HCT116 (Figure 1c), HEK293T and Jurkat cells (Supplementary Figure S1e). Coherently, we were able to pull down AGO2 by immunoprecipitating BAF155 (Figure 1d).

Since AGO1, another member of the AGO family, has a well-established role in transcriptional regulation of gene expression (9–11), we checked for an AGO1-SWI/SNF interaction, but our data do not support this hypothesis (Figure 1e).

SWI/SNF mainly exerts its function on chromatin, therefore we verified whether AGO2 is also associated with chromatin in our experimental model. We found that not only BAF155, but also AGO2 is associated with chromatin in HeLaS3 and HCT116 cell lines by a chromatin binding assay (Figure 2a), in agreement with previous reports (11). We next purified nuclei from HeLaS3 cells and verified that AGO2 and BAF155 interact in the nuclear compartment (Figure 2b).

Treatment with DNase I or RNase A before IP did not disrupt AGO2-BAF155 interaction (Figure 2c and d and Supplementary Figure S2), ruling out the possibility that DNA and/or RNA bridge the association of AGO2 with SWI/SNF.

Nuclear AGO2 is bound to a novel class of sRNAs

To better characterize the nuclear function of AGO2, we profiled sRNAs specifically bound to nuclear AGO2 in HeLaS3, Jurkat and HCT116 cells by massively parallel high throughput sequencing (Figure 3 and Supplementary Figure S3). To identify even low-abundant sRNA molecules we attained a sequencing depth of 100 million reads per sample (Supplementary Figure S4a), which is substantially higher than the depth employed for the same type of analysis reported elsewhere (41,42).

By iteratively mapping AGO2-bound sRNAs on miRNAs, tRNAs, rRNAs and Rfam (34) we identified a class of sRNAs, herein referred to as 'other sRNAs', which do not align to any previously annotated sRNA classes (miRNAs, tRNAs, snRNA, snoRNA etc; Supplementary Figure S3a and Figure S3a–c). A bioinformatics analysis using miRanalyzer (43) failed to identify putative novel miRNAs among AGO2-bound 'other sRNAs'. 'Other sRNAs' associated with AGO2 display a size distribution peaking at 22–24 nt, typical of AGO2-associated sRNAs (Figure 3b and Supplementary Figure S3d–f), whereas mock-IPed 'other sRNAs' have a heterogeneous size profile.

Typically, miRNA loci give rise to hundreds to thousands of sRNA reads with nearly identical 3′ and 5′ termini. To ascertain whether AGO2-associated 'other sRNAs' could contain novel miRNAs which escaped the miRanalyzer algorithm, we grouped 'other sRNAs' overlapping by at least one nt into clusters. Only 0.86% of all clusters are miRNA-like clusters. Four hundred one clusters (representing 27% of these miRNA-like clusters) lie within ±500 nt of TSS of protein coding genes, and might represent the human homologues of the 437 novel TSS-miRNAs reported in mouse ESC (42). On the contrary, the vast majority of clusters (99.14%) consists of <50 reads per cluster slightly differing in 3′ and 5′ termini (low copy clusters). Therefore, we focused all subsequent analyses on 'other sRNAs' falling into low copy clusters.

We compared the binding sites of SWI/SNF in HeLaS3 cells (20) with the genomic coordinates of 'other sRNAs'. We found that 'other sRNAs' mapping to SWI/SNF binding sites are enriched in AGO2-IP as compared to mock-IP (P-value $< 2.2 \times 10^{-16,}$ exact binomial test, Figure 3c). We speculated whether this new class of sRNAs bound to

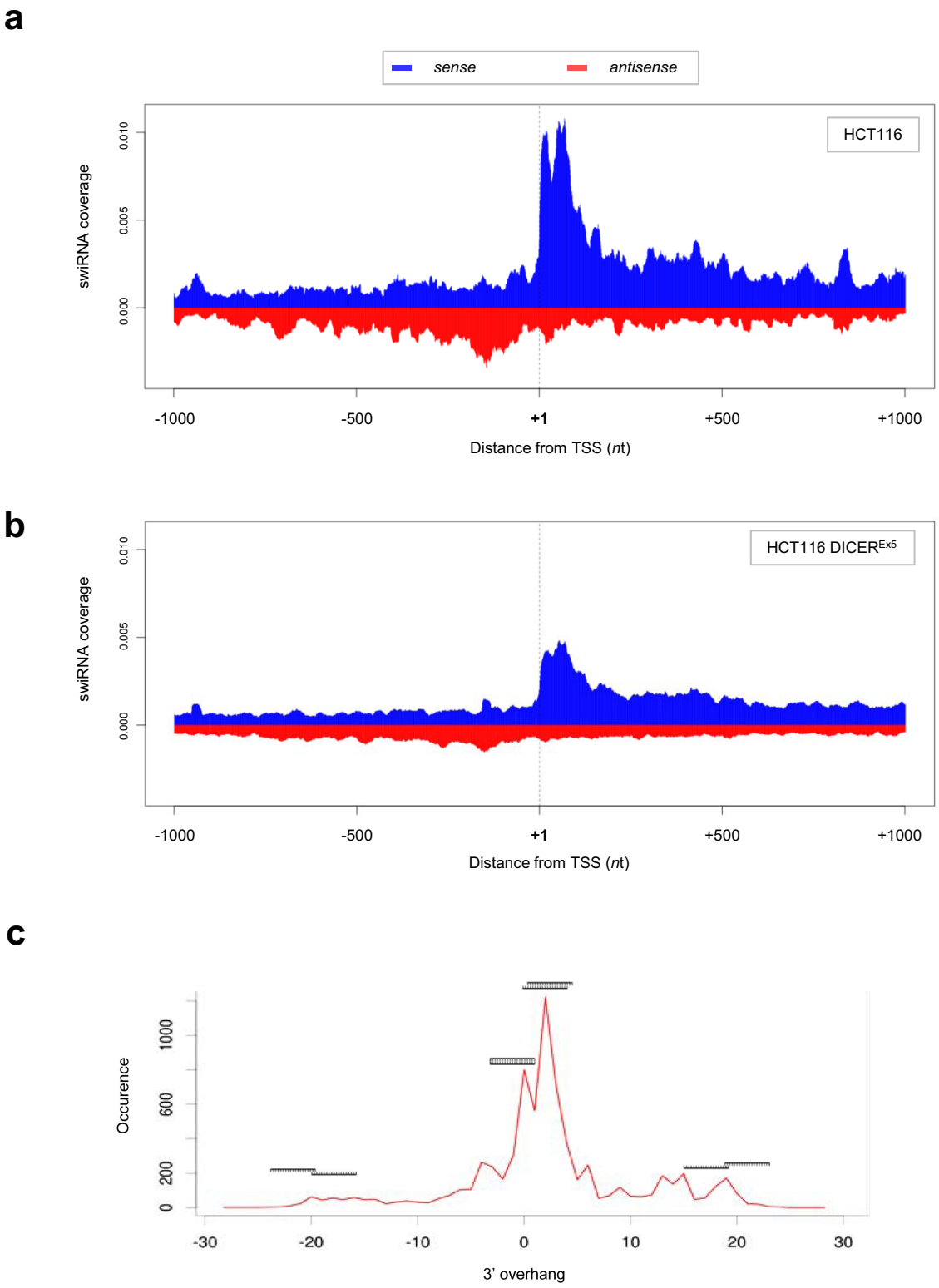

Figure 5. swiRNAs are processed in a Dicer-dependent manner.Coverage of swiRNAs in parental HCT116 cell line (**a**) and in HCT116 Dicer^{Ex5} cell line (**b**.) Coverage at each position represents the number of swiRNAs mapping at the indicated distance from a TSS, normalized by the total number of 'other sRNAs'. Only 'other sRNAs' lying within clusters of <50 molecules were considered. Red and blue represent sRNAs in the sense and anti-sense orientation with respect to gene transcription, respectively. (**c**) swiRNAs display a 2 nt 3′ overhang typical of Dicer-processed sRNAs. The length (nt) of 3′ overhang for pairs of complementary swiRNAs was plotted. Negative values represent a 5′ overhang. Most pairs display a 2 nt 3′ overhang, suggestive of Dicer processing.

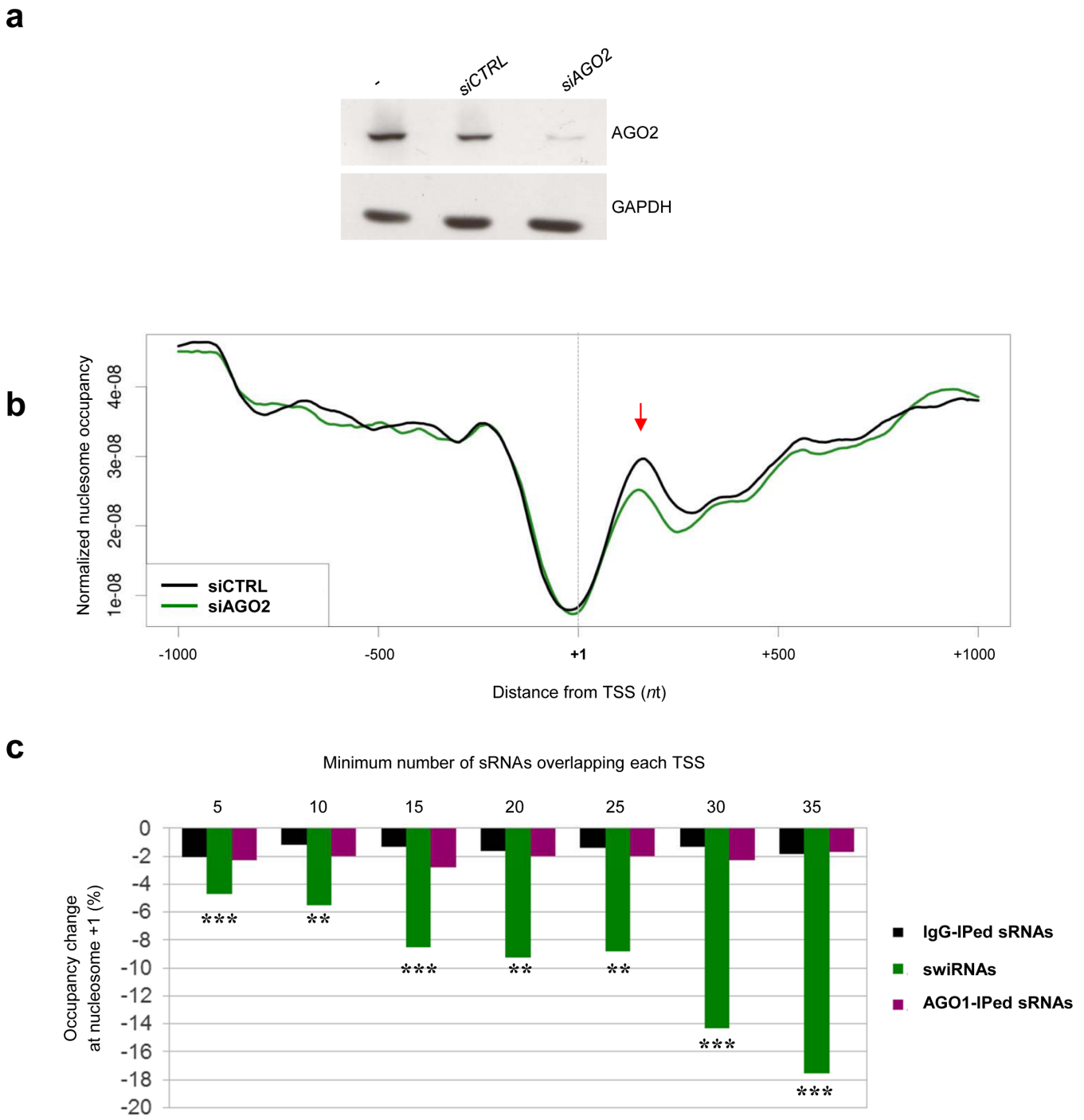

Figure 6. AGO2 knock-down affects nucleosome occupancy at TSSs bound by SWI/SNF. (**a**) HeLaS3 cells were transfected with a control siRNA (siCTRL) or a pool of AGO2 siRNA (siAGO2). Down-regulation of AGO2 protein was verified by western blot. GAPDH was used as loading control. (**b**) Chromatin from siCTRL- or siAGO2-treated HeLaS3 cells was digested by MNase and recovered DNA fragments were sequenced. Nucleosome occupancy profile for siCTRL and siAGO2 cells was plotted for TSSs with at least 30 swiRNAs (siCTRL, black line; siAGO2, green line). The occupancy at the nucleosome +1 (arrow) is reduced in AGO2 knock-down cells. (**c**) Bars height represents percent reduction of nucleosome occupancy (siAGO2 versus siCTRL) at TSS ±150 nt overlapped by at least the indicated number of swiRNAs (green), IgG-IP 'other sRNAs' (black) and AGO1-associated 'other sRNAs' (purple). $^{**}P$ value < 0.01, $^{***}P$ value $< 10^{-3}$, paired Wilcoxon test (see 'Materials and Methods' section for details).

AGO2 mapped nearby TSSs. In fact, it has been recently shown that in HeLaS3 cells SWI/SNF is recruited in the proximity of TSS of expressed genes (20). We found a specific enrichment of sRNAs mapping on TSS in AGO2-IP relative to mock-IP (P-value $< 2.2 \times 10^{-16}$, exact binomial test, Figure 3d and Supplementary Figure S3g–i). Interestingly, 'other sRNAs' bound by AGO2 preferentially map around TSSs bound by SWI/SNF (Figure 3e) as compared to TSSs not bound by SWI/SNF (Figure 3f). AGO2-bound 'other sRNAs' lie on both the sense and anti-sense direction with respect to gene transcription. These analyses clearly show that nuclear AGO2 is loaded with sRNAs arising from SWI/SNF bound TSSs. For this reason, we will herein refer to nuclear AGO2-associated sRNAs mapping within ±150 nt of a TSS as swiRNAs (Figure 4a). We found 4549 TSSs overlapped by at least one swiRNA. As expected, swiRNAs co-localize with histone modifications which are typical of open chromatin and TSS (H3K9ac, H2Az). However, while we observe a positive correlation between swiRNA abundance and the ChIP signal intensity of SWI/SNF components (Supplementary Figure S4b), such correlation is not observed for histone modifications (data not shown). Accordingly, although swiRNAs originate from transcriptionally active TSSs, there is no correlation between the number of swiRNAs overlapping a TSS and the expression level of that locus (Figure 4b). On the other hand, IgG IPed sRNAs mapping nearby TSSs correlate with the expression level of the corresponding gene. This observation indicates that swiRNA levels on a particular promoter are not under the control of the same mechanisms regulating transcription of that particular gene.

CLIP-seq approaches have recently been used to assess the specificity of protein–RNA interactions *in vivo*. Typically, UV-crosslinked RNA molecules recovered from CLIP experiments display single nucleotide substitutions, corresponding to nucleotide positions crosslinked to the protein of interest (44). We asked whether swiRNAs could be detected in published AGO2 CLIP datasets. We employed the same bioinformatics analysis described above on datasets SRR363673–6 (44) in which AGO2 CLIP was used to identify the interacting sRNAs. To the best of our knowledge, this is the only AGO2 CLIP dataset explicitly looking at sRNAs bound to AGO2, rather than at target mRNAs. Once we obtained the 'other sRNAs' in low copy clusters, we divided them into two classes: those perfectly mapping to human genome, and hence lacking the 'cross-linking signature' of CLIP (unCLIPPED), and those mapping with one or more mismatches (which are bona-fide AGO2-bound *in vivo*, CLIPPED). We found that 'CLIPPED sRNAs' represented the vast majority of the dataset (96.6% of 'other sRNAs in low copy clusters), corroborating the idea that AGO2 actually binds these molecules *in vivo*. Among CLIPPED 'other sRNAs' we found a proportion of swiRNAs (11.0%) even higher than the one observed in our HeLaS3 datasets (6.4%). Moreover, CLIPPED swiRNAs from this dataset displayed a size distribution compatible with AGO2 binding (data not shown). This analysis, although at much lower depth, recapitulates what we observed in our AGO2 and IgG RNA IP in HeLaS3, Jurkat and HCT116 cell lines, further confirming that swiRNAs are directly and specifically bound to AGO2 *in vivo*.

swiRNAs are processed by Dicer

In order to assess whether swiRNAs processing is Dicer dependent, we characterized by RNA-seq nuclear AGO2-bound sRNAs in parental HCT116 cells and in a Dicer hypomorphic subclone of HCT116 cells (22) (HCT116Dicer^{EX5}). We observed a drastic reduction of swiRNAs in HCT116Dicer^{EX5} relative to parental HCT116, indicating that Dicer processes swiRNAs (Figure 5a and b, Supplementary Figure S5). Dicer processes double stranded RNA (dsRNA) molecules which may result either by intramolecular pairing of an RNA molecule (e.g. hairpin structures) or by pairing of two complementary RNA molecules (e.g. pairing of RNAs transcribed from opposite strands of the same locus). Using RNAfold (45) we found that most swiRNAs are embedded in sequences potentially able to form secondary structures. However, analysis of the the predicted structures revealed that most (78%) swiRNAs do not lie within a miRNA-like hairpin, which is a prerequisite for DROSHA- and Dicer-mediated miRNA processing. The same algorithm correctly scored as miRNA-like structures most (77.4%) human miRNAs (miRbase v20).

We therefore explored whether swiRNAs may be processed from a dsRNA molecule arising from transcription of both strands of the same locus. Taking advantage of the HeLaS3 long RNA-seq dataset published by ENCODE, we looked for putative dsRNA precursors of swiRNAs. We found that ∼51% of the swiRNAs loci give rise to long RNA molecules derived from sense/antisense transcription. These RNA molecules embed the entire mature swiRNA and may form a dsRNA suitable for processing by Dicer. Since long noncoding RNA characterization in human cells is still incomplete, we do not exclude that a greater depth of sequencing could unveil dsRNA precursors for all swiRNAs. Dicer processes long dsRNAs producing short dsRNAs characterized by a typical 3′ overhang of two nucleotides. We found 2265 unique pairs of complementary swiRNAs mapping on opposite strands of the same locus and we looked at the 3′ termini of these pairs. Interestingly we observed a strong bias toward a 2 nt 3′ overhang (Figure 5c), further confirming that Dicer plays a pivotal role in swiRNA biogenesis.

Based on to the size, Dicer-mediated processing and association with AGO2, we conclude that swiRNAs are a novel class of sRNAs distinct from any other class of sRNAs mapping near TSSs (41–42,46–48).

AGO2 depletion affects nucleosome positioning in a swiRNA-dependent manner

The identification of novel interactions between AGO2, core components of SWI/SNF and swiRNAs suggests the involvement of AGO2 in SWI/SNF-mediated chromatin remodelling. Therefore, we depleted AGO2 in HeLaS3 cells by knock-down and performed MNase-Seq to assess nucleosome positioning around TSSs in control and AGO2 knock-down cells (Figure 6a).

Paired end sequencing of MNase-digested DNA fragments yielded ∼592 million reads in AGO2 knock-down and ∼711 million reads in control cells. The size distribution of DNA fragments is strikingly similar in the two samples,

ruling out the possibility of a bias due to unequal digestion (Supplementary Figure S6a–c). In our analysis we defined nucleosome occupancy as the frequency with which a given DNA region is occupied by a nucleosome within a cell population. We compared the average nucleosome occupancy profiles around all TSSs of expressed genes among the two samples; consistently with previous reports (49–51), we observed in both samples a nucleosome-free region located immediately upstream the TSS, flanked by two well-positioned nucleosomes, referred to as −1 and +1 nucleosome. We did not observe any changes in nucleosome phasing (data not shown). On the contrary we found a mild (1.98%) but highly significant (P-value $= 1.5 \times 10^{-15}$, paired t-test) decrease in the average occupancy at nucleosome +1 in AGO2 knockdown cells (data not shown). More interestingly, when we looked at TSSs from which at least 30 swiRNAs are transcribed, occupancy at nucleosome +1 was significantly affected with 14% decrease (paired Wilcoxon test P-value $=$ 0.0003723, FDR < 0.01) upon AGO2 depletion. This phenotype resembles the one observed in murine cells where core components of SWI/SNF are genetically ablated (21), suggesting that AGO2 and SWI/SNF not only physically interact but also functionally cooperate in determining nucleosome occupancy at TSS.

Moreover, AGO2-dependent reduction in nucleosome occupancy positively correlated with the number of swiRNAs mapping within ±150 nt from TSS (Figure 6b). As a control, we performed the same analysis using IgG-IP sRNAs for TSS selection (Figure 6b). In agreement with the fact that AGO1 does not interact with SWI/SNF (Figure 1e), no correlation is observed between the number of AGO1-associated sRNAs overlapping with each TSS and occupancy reduction at nucleosome +1 in AGO2 knockdown cells (Figure 6c; Supplementary Figure S6d).

A biological process GO term analysis of all genes expressed in HeLaS3 sorted by decreasing number of swiRNAs sheds light on the function of genes whose nucleosome occupancy is affected by AGO2 depletion (Supplementary Table S3). Top scoring enriched GO terms pinpoint at nucleic acid and protein metabolism.

These data suggest that swiRNA, AGO2 and SWI/SNF collaborate to determine nucleosome occupancy at TSS of genes involved in key metabolic pathways.

DISCUSSION

By a SILAC-based quantitative proteomic analysis, we characterized the endogenous AGO2 interactome. In interaction studies quantitative proteomic approaches have the potential to easily distinguish true protein:protein interactions from experimental contaminants by allowing the accurate relative quantification of protein abundance in both control and test immunoprecipitation samples (52–54). Taking advantage of this powerful approach, we describe a physical and functional interplay between components of the RNAi and chromatin remodelling machineries in human cells for the first time. Indeed, we show that nuclear AGO2 interacts with SWI/SNF, and a new class of sRNAs that we termed swiRNAs. These novel endogenous, Dicer-dependent sRNAs arise from SWI/SNF bound TSSs and determine nucleosome occupancy on target TSSs, likely by recruiting nuclear AGO2–SWI/SNF complexes onto specific TSSs.

swiRNAs do not overlap with any of the reported classes of sRNAs and therefore represent a novel class of sRNAs. In fact, unlike tiRNAs (48) and TSSa RNAs (41,46), swiRNAs have a peak size of 22–24 nt, are processed from long dsRNA precursors by Dicer and are associated with AGO2 in the nuclear compartment. Interestingly swiRNAs are directly bound to AGO2 *in vivo*, as shown by AGO2 CLIP datasets analysis, demonstrating that swiRNAs are not a mere by-product of transcription. This is in agreement with the finding that gene expression level does not correlate to the number of swiRNAs mapping nearby each TSS. Therefore, their abundance and stability might be regulated by specific molecular mechanisms in response to intra- and/or extra-cellular signals.

Based on our sequencing data, we estimate that swiRNAs are on average 10 000-fold less abundant than miRNAs. However, a miRNA has about a hundred of target mRNA species, often with multiple binding sites for each target. On the contrary, a swiRNA has a single possible genomic target, present in two copies per cell. Indeed, even a single swiRNA molecule per TSS would suffice to drive recruitment of AGO2–SWI/SNF and affect nucleosome positioning, as shown by our data. It is worth mentioning that swiRNAs could be selectively expressed in specific phases of cell cycle. Moreover, they might be 5′ and/or 3′ modified and our standard cloning conditions which are optimized for unmodified small RNA molecules could underestimate swiRNA abundance.

In eukaryotes, nucleosome occupancy around TSS of actively transcribed genes is characterized by a typical pattern consisting of a nucleosome free region immediately upstream of TSS and a very well positioned +1 nucleosome immediately downstream of TSS. It has been recently shown that genetic ablation of key subunits of SWI/SNF in murine cells specifically reduces nucleosome occupancy at TSS, suggesting that SWI/SNF function is especially relevant for the proper positioning of nucleosomes around TSS (21). Our data highlight that depletion of AGO2 in human cells decreases nucleosome +1 occupancy and this positively correlates with the number of swiRNAs mapping around a particular TSS. These observations support the idea that in mammalian cells SWI/SNF complex specifically controls nucleosome +1 occupancy in cooperation with AGO2 and swiRNAs.

The widespread conservation across eukaryotes of a typical nucleosome occupancy pattern around actively transcribed TSSs (49) is suggestive of a functional role for such pattern. Previous reports highlight that occupancy at nucleosome +1 is not predictive of gene expression level (55). Accordingly, transcriptome analysis upon depletion of AGO2 failed to reveal any significant change in steady-state expression level of the genes displaying higher numbers of swiRNAs, or any other apparent alteration of mRNA expression (data not shown). Although several hypotheses have been proposed, the exact functional role of nucleosome occupancy at nucleosome +1 still remains elusive. It has been previously suggested that nucleosome occupancy pattern in promoters might be related to plasticity in gene expression (56). Prompted by our observation

that swiRNA-enriched genes are mainly involved in protein and nucleic acid metabolisms, we speculate that such genes might share a common mechanism to fine tune their expression in response to environmental and cellular conditions (i.e. cell cycle progression, growth signals or metabolic stress). Proper occupancy at nucleosome +1 could play a role in the response to stimuli or changes in environmental conditions that require a rapid and prompt adaptation of cell metabolic rates. Further investigation of the cellular outcome of AGO2 depletion might shed light on the physiological consequences of altered occupancy at nucleosome +1.

We propose that swiRNAs mediate recruitment of AGO2 and SWI/SNF onto targeted TSS, thus contributing to proper positioning of nucleosome +1. This is the first demonstration that the involvement of RNAi in chromatin remodelling is conserved not only in plants and fungi but also in mammals. Interestingly, both AGO2 and SWI/SNF have been previously independently reported to be involved in DNA damage response (15,54) and in alternative splicing (13,56). Our data are suggestive of a functional interplay of AGO2 and SWI/SNF in these processes and represent an important starting point for further investigations.

ACCESSION NUMBERS

Sequencing data has been deposited in the ArrayExpress Database under accession numbers: E-MTAB-1844; E-MTAB-2056; E-MTAB-2084; E-MTAB-2104; E-MTAB-2129; E-MTAB-2128.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Leandro Castellano for providing HCT116 and HCT116 Dicer^{EX5} cell lines, Dr Carlo Cogoni and Dr Cecilia Poli for critical reading of the manuscript and Katia Broccoletti for laboratory support.

FUNDING

Epigenomic Flagship Project (EPIGEN) MIUR-CNR (to V.F.); Associazione Italiana per la Ricerca sul Cancro; Giovanni Armenise-Harvard Foundation Career Development; Association of International Cancer Research; Italian Ministry of Health (to T.B.). Umberto Veronesi Foundation (to M.M.). Funding for open access charge: Epigenomic Flagship Project (Epigenomica EPIGEN) MIUR-CNR (to V.F.).
Conflict of interest statement. None declared.

REFERENCES

1. Napoli,C., Lemieux,C. and Jorgensen,R. (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell*, **2**, 279–289.
2. Romano,N. and Macino,G. (1992) Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences. *Mol. Microbiol.*, **6**, 3343–3353.
3. Tabara,H., Sarkissian,M., Kelly,W.G., Fleenor,J., Grishok,A., Timmons,L., Fire,A. and Mello,C.C. (1999) The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell*, **99**, 123–132.
4. Fagard,M., Boutet,S., Morel,J.B., Bellini,C. and Vaucheret,H. (2000) AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11650–11654.
5. Meister,G. (2013) Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.*, **14**, 447–459.
6. Grewal,S.I.S. and Jia,S. (2007) Heterochromatin revisited. *Nat. Rev. Genet.*, **8**, 35–46.
7. Matzke,M., Kanno,T., Daxinger,L., Huettel,B. and Matzke,A.J.M. (2009) RNA-mediated chromatin-based silencing in plants. *Curr. Opin. Cell Biol.*, **21**, 367–376.
8. Morris,K.V., Santoso,S., Turner,A.-M., Pastori,C. and Hawkins,P.G. (2008) Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. *PLoS Genet.*, **4**, e1000258.
9. Kim,D.H., Villeneuve,L.M., Morris,K.V. and Rossi,J.J. (2006) Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat. Struct. Mol. Biol.*, **13**, 793–797.
10. Chu,Y., Yue,X., Younger,S.T., Janowski,B.A. and Corey,D.R. (2010) Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. *Nucleic Acids Res.*, **38**, 7736–7748.
11. Huang,V., Zheng,J., Qi,Z., Wang,J., Place,R.F., Yu,J., Li,H. and Li,L.-C. (2013) Ago1 interacts with RNA polymerase II and binds to the promoters of actively transcribed genes in human cancer cells. *PLoS Genet.*, **9**, e1003821.
12. Alló,M., Buggiano,V., Fededa,J.P., Petrillo,E., Schor,I., de la Mata,M., Agirre,E., Plass,M., Eyras,E., Elela,S.A. *et al.* (2009) Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.*, **16**, 717–724.
13. Ameyar-Zazoua,M., Rachez,C., Souidi,M., Robin,P., Fritsch,L., Young,R., Morozova,N., Fenouil,R., Descostes,N., Andrau,J.-C. *et al.* (2012) Argonaute proteins couple chromatin silencing to alternative splicing. *Nat. Struct. Mol. Biol.*, **19**, 998–1004.
14. Wei,W., Ba,Z., Gao,M., Wu,Y., Ma,Y., Amiard,S., White,C.I., Rendtlew Danielsen,J.M., Yang,Y.-G. and Qi,Y. (2012) A role for small RNAs in DNA double-strand break repair. *Cell*, **149**, 101–112.
15. Gao,M., Wei,W., Li,M.-M., Wu,Y.-S., Ba,Z., Jin,K.-X., Li,M.-M., Liao,Y.-Q., Adhikari,S., Chong,Z. *et al.* (2014) Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. *Cell Res.*, **24**, 532–541.
16. Euskirchen,G., Auerbach,R.K. and Snyder,M. (2012) SWI/SNF chromatin-remodeling factors: multiscale analyses and diverse functions. *J. Biol. Chem.*, **287**, 30897–30905.
17. Clapier,C.R. and Cairns,B.R. (2009) The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.*, **78**, 273–304.
18. Ho,L., Jothi,R., Ronan,J.L., Cui,K., Zhao,K. and Crabtree,G.R. (2009) An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 5187–5191.
19. De,S., Wurster,A.L., Precht,P., Wood,W.H. 3rd, Becker,K.G. and Pazin,M.J. (2011) Dynamic BRG1 recruitment during T helper differentiation and activation reveals distal regulatory elements. *Mol. Cell. Biol.*, **31**, 1512–1527.
20. Euskirchen,G.M., Auerbach,R.K., Davidov,E., Gianoulis,T.A., Zhong,G., Rozowsky,J., Bhardwaj,N., Gerstein,M.B. and Snyder,M. (2011) Diverse roles and interactions of the SWI/SNF chromatin remodeling Complex revealed using global approaches. *PLoS Genet*, **7**, e1002008.
21. Tolstorukov,M.Y., Sansam,C.G., Lu,P., Koellhoffer,E.C., Helming,K.C., Alver,B.H., Tillman,E.J., Evans,J.A., Wilson,B.G., Park,P.J. *et al.* (2013) Swi/Snf chromatin remodeling/tumor suppressor complex establishes nucleosome occupancy at target promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 10165–10170.
22. Cummins,J.M., He,Y., Leary,R.J., Pagliarini,R., Diaz,L.A., Sjoblom,T., Barad,O., Bentwich,Z., Szafranska,A.E., Labourier,E. *et al.* (2006) The colorectal microRNAome. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 3687–3692.

23. Meister,G., Landthaler,M., Patkaniowska,A., Dorsett,Y., Teng,G. and Tuschl,T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell*, **15**, 185–197.
24. Cernilogar,F.M., Onorati,M.C., Kothe,G.O., Burroughs,A.M., Parsi,K.M., Breiling,A., Lo Sardo,F., Saxena,A., Miyoshi,K., Siomi,H. *et al.* (2011) Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. *Nature*, **480**, 391–395.
25. Shevchenko,A., Tomas,H., Havlis,J., Olsen,J.V. and Mann,M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.*, **1**, 2856–2860.
26. Rappsilber,J., Mann,M. and Ishihama,Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.*, **2**, 1896–1906.
27. Cox,J. and Mann,M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.*, **26**, 1367–1372.
28. Cox,J., Neuhauser,N., Michalski,A., Scheltema,R.A., Olsen,J.V. and Mann,M. (2011) Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.*, **10**, 1794–1805.
29. Kim,D., Pertea,G., Trapnell,C., Pimentel,H., Kelley,R. and Salzberg,S.L. (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.*, **14**, R36.
30. Trapnell,C., Hendrickson,D.G., Sauvageau,M., Goff,L., Rinn,J.L. and Pachter,L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.*, **31**, 46–53.
31. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, **17**, pp. 10–12.
32. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
33. Kozomara,A. and Griffiths-Jones,S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, **39**, D152–D157.
34. Burge,S.W., Daub,J., Eberhardt,R., Tate,J., Barquist,L., Nawrocki,E.P., Eddy,S.R., Gardner,P.P. and Bateman,A. (2012) Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res.*, doi:10.1093/nar/gks1005.
35. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
36. Flicek,P., Ahmed,I., Amode,M.R., Barrell,D., Beal,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fairley,S. *et al.* (2012) Ensembl 2013. *Nucleic Acids Res.*, **41**, D48–D55.
37. Rosenbloom,K.R., Sloan,C.A., Malladi,V.S., Dreszer,T.R., Learned,K., Kirkup,V.M., Wong,M.C., Maddren,M., Fang,R., Heitner,S.G. *et al.* (2013) ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.*, **41**, D56–D63.
38. Eden,E., Navon,R., Steinfeld,I., Lipson,D. and Yakhini,Z. (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, **10**, 48.
39. Pfaff,J., Hennig,J., Herzog,F., Aebersold,R., Sattler,M., Niessing,D. and Meister,G. (2013) Structural features of Argonaute-GW182 protein interactions. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E3770–3779.
40. Wu,C., So,J., Davis-Dusenbery,B.N., Qi,H.H., Bloch,D.B., Shi,Y., Lagna,G. and Hata,A. (2011) Hypoxia potentiates microRNA-mediated gene silencing through posttranslational modification of Argonaute2. *Mol. Cell. Biol.*, **31**, 4760–4774.
41. Valen,E., Preker,P., Andersen,P.R., Zhao,X., Chen,Y., Ender,C., Dueck,A., Meister,G., Sandelin,A. and Jensen,T.H. (2011) Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. *Nat. Struct. Mol. Biol.*, **18**, 1075–1082.
42. Zamudio,J.R., Kelly,T.J. and Sharp,P.A. (2014) Argonaute-bound small RNAs from promoter-proximal RNA polymerase II. *Cell*, **156**, 920–934.
43. Hackenberg,M., Rodriguez-Ezpeleta,N. and Aransay,A.M. (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. *Nucleic Acids Res.*, **39**, W132–W138.
44. Stark,T.J., Arnold,J.D., Spector,D.H. and Yeo,G.W. (2012) High-resolution profiling and analysis of viral and host small RNAs during human cytomegalovirus infection. *J. Virol.*, **86**, 226–235.
45. Lorenz,R., Bernhart,S.H., Höner Zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol. AMB*, **6**, 26.
46. Seila,A.C., Calabrese,J.M., Levine,S.S., Yeo,G.W., Rahl,P.B., Flynn,R.A., Young,R.A. and Sharp,P.A. (2008) Divergent transcription from active promoters. *Science*, **322**, 1849–1851.
47. Core,L.J., Waterfall,J.J. and Lis,J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science*, **322**, 1845–1848.
48. Taft,R.J., Glazov,E.A., Cloonan,N., Simons,C., Stephen,S., Faulkner,G.J., Lassmann,T., Forrest,A.R.R., Grimmond,S.M., Schroder,K. *et al.* (2009) Tiny RNAs associated with transcription start sites in animals. *Nat. Genet.*, **41**, 572–578.
49. Schones,D.E., Cui,K., Cuddapah,S., Roh,T.-Y., Barski,A., Wang,Z., Wei,G. and Zhao,K. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell*, **132**, 887–898.
50. Jiang,C. and Pugh,B.F. (2009) Nucleosome positioning and gene regulation: advances through genomics. *Nat. Rev. Genet.*, **10**, 161–172.
51. Hartley,P.D. and Madhani,H.D. (2009) Mechanisms that specify promoter nucleosome location and identity. *Cell*, **137**, 445–458.
52. Ranish,J.A., Yi,E.C., Leslie,D.M., Purvine,S.O., Goodlett,D.R., Eng,J. and Aebersold,R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.*, **33**, 349–355.
53. Mann,M. (2006) Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.*, **7**, 952–958.
54. Vermeulen,M., Hubner,N.C. and Mann,M. (2008) High confidence determination of specific protein–protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.*, **19**, 331–337.
55. Huebert,D.J., Kuan,P.-F., Keleş,S. and Gasch,A.P. (2012) Dynamic changes in nucleosome occupancy are not predictive of gene expression dynamics but are linked to transcription and chromatin regulators. *Mol. Cell. Biol.*, **32**, 1645–1653.
56. Tirosh,I. and Barkai,N. (2008) Two strategies for gene regulation by promoter nucleosomes. *Genome Res.*, **18**, 1084–1091.