Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion

Cameron P Bracken¹,², Xiaochun Li¹, Josephine A Wright¹, David M Lawrence³, Katherine A Pillman¹, Marika Salmanidis³, Matthew A Anderson¹, B Kate Dredge³, Philip A Gregory¹,², Anna Tsykin¹, Corine Neilsen³, Daniel W Thomson³, Andrew G Bert³, Joanne M Leerberg⁴, Alpha S Yap⁵, Kirk B Jensen³, Yeesim Khew-Goodall¹,³,*,† & Gregory J Goodall¹,²,³,**,†

Abstract

The microRNAs of the miR-200 family maintain the central characteristics of epithelia and inhibit tumor cell motility and invasiveness. Using the Ago-HITS-CLIP technology for transcriptome-wide identification of direct microRNA targets in living cells, along with extensive validation to verify the reliability of the approach, we have identified hundreds of miR-200a and miR-200b targets, providing insights into general features of microRNA target site selection. Gene ontology analysis revealed a predominant effect of the miR-200 targets in widespread coordinate control of actin cytoskeleton dynamics. Functional characterization of the miR-200 targets indicates that they constitute subnetworks that underlie the ability of cancer cells to migrate and invade, including coordinate effects on Rho-ROCK signaling, invadopodia formation, MMP activity, and focal adhesions. Thus, the miR-200 family maintains the central characteristics of the epithelial phenotype by acting on numerous targets at multiple levels, encompassing both cytoskeletal effectors that control actin filament organization and dynamics, and upstream signals that locally regulate the cytoskeleton to maintain cell morphology and prevent cell migration.

Keywords cytoskeleton; HITS-CLIP; invadopodia; microRNA; miR-200

Subject Categories Cancer; Cell Adhesion, Polarity & Cytoskeleton; RNA Biology

DOI 10.15252/embj.201488641 | Received 2 April 2014 | Revised 6 June 2014 | Accepted 12 June 2014 | Published online 28 July 2014

The EMBO Journal (2014) 33: 2040–2056

Introduction

MicroRNAs and their targets are important components in the regulatory networks that maintain cell phenotype and control cell differentiation. Although microRNAs typically act as mild modulators of gene expression, exerting only a modest inhibitory effect on individual targets, conceivably they can broadly refine gene expression patterns because each microRNA may target several hundred different mRNAs. Thus, one microRNA can potentially influence a biological process by having a coordinated effect on multiple components of a network or pathway. However, due to the uncertainties in predicting or experimentally identifying the spectrum of targets of individual microRNAs, there are few confirmed examples of broad network regulation by a microRNA.

The miR-200 family of microRNAs acts as enforcers of the epithelial phenotype. They are expressed in most, if not all, epithelial cells and their expression must be turned off for epithelial to mesenchymal transition (EMT) to occur (Burk et al., 2008; Gregory et al., 2008a; Korpal et al., 2008; Park et al., 2008). EMT involves a morphological change whereby immobile epithelial cells acquire pro-invasive mesenchymal characteristics that are key to various developmental processes and are drivers of metastatic progression in cancer (Nieto, 2013). Two major targets of miR-200 in controlling EMT are the transcription repressors, ZEB1 and ZEB2 (formerly known as SIP1) (Gregory et al., 2008a; Park et al., 2008). Furthermore, the ZEB proteins are strong repressors of transcription of the miR-200 genes, producing a double negative feedback loop that is central to the control of EMT in in vitro models (Bracken et al., 2008; Burk et al., 2008; Parker et al., 2009; Gregory et al., 2011) and has been implicated in contributing to cancer progression through promotion of EMT in various epithelial-derived cancers. However, like all microRNAs, the miR-200 family members are predicted to...
have many other targets. Some of these may also play central roles in mediating the epithelial response to miR-200, while others may provide a more subtle role in the sculpting of phenotype by miR-200.

The miR-200 family consists of 5 miRNAs that are closely related in sequence, but are predicted to comprise two distinct classes in terms of their targets, with miR-200a and miR-141 sharing identical seed regions and miR-200b, miR-200c, and miR-429 constituting a separate targeting class. These 5 microRNAs arise from two precursor genes encoded at separate genomic loci. The expression of the 5 family members is highly correlated, being coexpressed in essentially all epithelial cells and absent in mesenchymal cells. While the targeting of ZEB1 and ZEB2 by miR-200 has been extensively investigated, and several additional targets have been identified, a comprehensive unbiased investigation of direct target genes, and the consequences for cell phenotype of the targeting of these genes, has not been reported.

MicroRNAs function as the specificity component of the protein–RNA complex known as RISC (RNA-induced silencing complex). The microRNA provides sequence-specific binding of RISC to specific mRNA targets, resulting in decreased efficiency of translation and an increased rate of mRNA degradation (Carthew & Sontheimer, 2009). While microRNAs are now relatively easy to discover and measure, the key to understanding their functions remains the identification of their gene targets, which until recently has largely been achieved via computational prediction followed by individual experimental verification, or miRNA manipulation in conjunction with approaches such as microarray or proteomic profiling. In silico target prediction is limited by our incomplete understanding of targeting ‘rules’ due largely to an inability to reliably model the influences of RNA secondary structure and RNA-binding proteins that interfere with potential target sites. Approaches based on mRNA expression analysis can only identify targets that are destabilized at the RNA level, cannot identify the precise site of targeting, and are unable to differentiate direct from indirect targets, while proteomic approaches are limited in their sensitivity and also do not differentiate direct from indirect targets.

A considerable methodological improvement has been the development of the Ago-HITS-CLIP (Argonaute High Throughput Sequencing after Cross-Linked Immunoprecipitation) procedure, in which RNA–protein complexes are stabilized by UV cross-linking in live cells, followed by direct immunoprecipitation and purification of miRNA-loaded RISC, enabling the identification of directly associated target transcripts on a global scale by massively parallel sequencing (Chi et al, 2009). By capturing the RISC complex in the act of binding to mRNA in living cells, and incorporating several purification steps, Ago-HITS-CLIP has the advantages of being highly stringent, requires no a priori predictions of the identity or locations of binding sites, and avoids non-specific Ago–RNA interactions that may otherwise occur in vitro (Riley et al, 2012). We have applied this procedure to identify many transcripts that interact with miR-200, including a number of non-canonical targets such as central-paired and seed-mismatch interactions, and also find novel target sites in previously misannotated transcripts. We find that regulators of actin cytoskeleton dynamics are strongly enriched among the targets of both miRNAs, indicating that the miR-200 family imposes coordinated control of functional networks that are central to cell shape and motility, and of crucial importance in cancer cell invasion and metastasis. In particular, we show that the formation of invadopodia, which relies on rearrangement of the actin cytoskeleton and provides a site for localized secretion of enzymes to degrade the extracellular matrix, is regulated by miR-200 at multiple points in the pathway.

Results

Global identification of transcripts targeted by miR-200

The 5 members of the miR-200 family are coregulated and have a controlling influence on the epithelial versus mesenchymal cell state. They are closely related to each other in sequence, but fall into 2 classes with respect to their seed regions, indicating they are likely to have two distinct sets of mRNA targets. We investigated miR-200a as representative of the miR-200a/miR-141 seed class and miR-200b as representative of the miR-200b/miR-200c/miR-429 seed class. We performed Ago-HITS-CLIP on control and transfectected MDA-MB-231 breast cancer cells, which have very low endogenous levels of the miR-200 family (Gregory et al, 2008a). We confirmed the successful precipitation of Ago by immunoblot analysis (Supplementary Fig S1A) and subjected the barcoded cDNAs (Supplementary Fig S1B) to Illumina sequencing. From the mapped small RNA reads, 223 unique miRNAs were identified with miR-27a being the most abundant. Upon transfection, Ago-bound miR-200a and miR-200b were each increased to levels similar to miR-27a, confirming the loading of Ago with the transfected miRNAs at physiological levels (Supplementary Fig S1C).

From the pool of immunoprecipitated miRNAs, reads were mapped across 12,814 genes. Sites bound to Ago were identified by clustering overlapping reads into peaks, setting a threshold of at least 3 non-identical overlapping reads for designation as a peak. Although 43% of the individual reads mapped to introns and intergenic regions, this drops to 14% after the clustering of individual reads into overlapping peaks. Conversely, reads mapping to 3'UTRs and coding regions increase from 55% of individual reads to 83% of peaks, suggesting the clustering of reads into peaks reduces background sequencing noise (Fig 1A) and identifies prominent interaction hotspots. To identify transcripts targeted by miR-200, we located read peaks from cells transfected with miR-200a or miR-200b that were absent in samples from untransfected cells and from cells transfected with a control miRNA mimic. Because the majority of interactions occur through base pairing to the miRNA ‘seed region’, we initially identified such interaction sites by cross referencing peaks that were dependent on miR-200 transfection with all the potential cognate 6-, 7-, or 8-mer seed sites. This identified 917 and 1,194 transcripts that directly bound miR-200a and miR-200b, respectively (Supplementary Table S1). Of these, ~65% map to 3'UTRs, 33% to coding regions, and 2% to 5'UTRs, with 9% representing 8-mers, 27% 7-mers and the remaining 64% 6-mers. Many potential seed matches were not sites of interaction, highlighting the utility of Ago-HITS-CLIP in distinguishing functional from predicted sites. For example, only a small subset of the potential seed sites in the CFL2 and MPRIp mRNAs are actually engaged by miR-200 at a discernable level (Fig 1B and C).

In addition to canonical target interactions whereby the miRNA seed region perfectly base pairs with a target transcript, miRNAs
sometimes interact through seed sites that contain mismatches or bulges (Chi et al., 2012; Loeb et al., 2012) or via ‘central-pairing’ which involves the central region of the miRNA but not the seed region (Shin et al., 2010). We found examples of seed-bulge and central-paired interactions with miR-200, but these were rare, cumulatively accounting for only ~3% of target sites. In contrast, peaks attributable to seed mismatches were abundant, constituting ~35% of all interaction sites within 3’UTRs. Examples of various non-canonical miR-200 interactions are shown in Supplementary Fig S2. Because the miR-200a and miR-200b seed regions differ at one position, they are each equivalent to a seed mismatch of the other. We found that 15% of seed-based interaction sites for each miR was also a site of interaction for the other family member, demonstrating there is a high degree of specificity in seed region interactions, but confirming that mismatch interactions can occur at a proportion of sites.

Targeting by miRNAs typically reduces the level of the target mRNA (Grimson et al., 2007; Hafner et al., 2010). By microarray analysis, we found a strong bias toward downregulation by miR-200 of mRNAs with miR-200 interaction sites identified by Ago-HITS-CLIP, with the greatest bias being for transcripts containing perfect 8-mer seed sites within 3’UTRs (Supplementary Table S1), consistent with previous reports that longer seed matches correlate with stronger target repression (Grimson et al., 2007). These biases were also present, but progressively less strong, for 7-mer and 6-mer sites. Transcripts targeted through seed sites in the coding region were less strongly biased toward downregulation, in agreement with studies of let-7, miR-1, and miR-124 (Lim et al., 2005; Forman & Coller, 2010). Pronounced RNA destabilization via seed-mismatch sites was rare, indicating that although seed-mismatch sites are relatively frequent sites of miRNA association, they rarely have a pronounced effect on mRNA abundance. This observation, coupled with the rarity of other non-canonical miRNA interaction sites, led us to focus on the effects of miR-200 on genes targeted through seed matches in 3’UTRs.

The HITS-CLIP procedure indicates sites of interaction between Ago and mRNA, but additional data are needed to determine the strength of repression that this interaction confers on the mRNA. To better assess the functional effects of the miRNA-target interactions identified by the Ago-HITS-CLIP, we selected 56 genes that represent a range of perfect seed matches (6-mer, 7-mer, and 8-mer) and seed-mismatch sites for further verification. Given the reported predominance of 3’UTR targeting and our own data showing the expression of genes possessing 3’UTR sites are more likely to be downregulated by miRNAs (Supplementary Table S1), we restricted our selection of genes to those with interaction sites within their 3’UTRs and tested the capacity of miR-200 to regulate these sites by the cloning of full-length 3’UTRs into luciferase reporter constructs. Full-length 3’UTRs, rather than isolated minimal target sites, were used in order to more fully replicate the endogenous site of targeting. The luciferase reporters confirmed that the majority of targets identified by Ago-HITS-CLIP are indeed repressed by miR-200 (Fig 2A) and, as indicated by our microarray data (Supplementary Table S1), demonstrated that repression is particularly strong when mediated through perfect 8-mer seed sites (Fig 2C). For example, we found miR-200b
confers statistically significant repression of 84% of the 8-mer and 7-mer sites selected, with most of these targets (81%) being strongly repressed to a level < 75% of the control. Of the 6-mer sites tested, 48% were significantly repressed, although only 13% of these were suppressed by more than 75%. The same hierarchical effect (8-mer > 7-mer > 6-mer seed matches) on endogenous mRNA levels was also observed by qPCR measurement of the effect of miR-200 on the endogenous mRNAs (Fig 2B and D). The targeting of a subset of genes possessing 8-mer miR-200 seed sites, selected on the basis of their important roles in cytoskeletal remodeling, was further confirmed by Western blot analysis (Fig 2E). Although individual targets may be repressed through seed-mismatch sites, in neither the luciferase reporters, qPCR or microarray data were seed mismatches found to generally mediate functionally significant repression.

The unbiased nature of Ago-HITS-CLIP allows the identification of interaction sites that are missed by prediction tools due to misannotation of transcripts. For example, we found that a prominent miR-200b target site resides downstream of the annotated end of the EGFR mRNA (Fig 3A). Inspection of ESTs aligned to this region supports the existence of a long UTR form. We designed qPCR assays to quantitate expression of the short and long forms, and found that the long form is predominant in MDA-MB-231 cells and is significantly reduced in level in response to miR-200b (Fig 3B). We cloned the long and short UTRs into luciferase reporters and verified that miR-200b significantly repressed activity of the reporter with the long 3’UTR, but not the short form (Fig 3C). This suggests that EGFR expression, which can have important consequences in tumors, can be subjected to regulation in response to miR-200, but can made independent of miR-200 by alternative polyadenylation.

**HITS-CLIP reveals a network of cytoskeletal regulators, including components of the Rho signaling pathway, that are miR-200 target genes**

To assess whether any functional systems might be coordinately regulated by miR-200, we conducted gene ontology analysis of the miR-200 HITS-CLIP targets. This revealed a striking over-representation of genes associated with cytoskeletal remodeling (Fig 4A). Importantly, cytoskeletal remodeling was the pathway most highly targeted by both miR-200a and miR-200b, and this was even more pronounced when restricted to genes with binding sites within their 3’UTRs (Supplementary Fig S3), the most functionally optimal site of targeting. Most of the other significantly enriched ontologies are the functional outcomes of cytoskeletal remodeling, such as cell adhesion, and processes with which miR-200 has been previously implicated including EMT, TGF-β, and Wnt signaling.

A number of the miR-200 targets identified by Ago-HITS-CLIP are upstream regulators or downstream effectors of the Rho-ROCK signaling pathway, which is implicated in tumor progression and invasion (Sahai & Marshall, 2003; Croft et al, 2004; Fig 4B). RhoA itself was not a strong direct target, but the RhoA activating guanine exchange factors (GEFs), ARHGEF3 and NET1 (ARHGEF8), and the inactivating GTPase-activating protein (GAP), ARHGAP29, were identified as targets (Fig 2). In light of their opposing roles on Rho activity, we determined the net effect of miR-200 by measuring the levels of both active and total Rho following transfection with miR-200a, miR-200b, or both. MiR-200b decreased the levels of active Rho without affecting total Rho levels (Fig 4C), consistent with its targeting of the Rho activators ARHGEF3 and NET1. In addition, the concomitant direct downregulation of ROCK2 levels by miR-200 (Fig 2) might amplify its effect on RhoA-ROCK signaling. In contrast, we did not observe any significant effect of miR-200 on the activation of cdc42 or Rac1 (data not shown).

**MiR-200 and its targets regulate cell morphology, migration, and invasion**

As indicated by gene ontology analysis, the pathway with the largest cohort of miR-200 targets is that associated with cytoskeletal remodeling which is consistent with previous work implicating miR-200 in controlling cell motility and cell shape changes (Burk et al, 2008; Gregory et al, 2008a). We also verified that, in breast cancer cells, miR-200 promoted the rearrangement of the actin cytoskeleton from stress fibers to cortical actin, (Supplementary Fig S4A) and inhibited cell migration and invasion (Supplementary Fig S4B-D). We then used a real-time invasion assay to compare the effect of knockdown of several individual miR-200 targets with the effect of miR-200 itself. Knockdown of MPRIP, ABL2, and WIPF1 each inhibited invasion (Fig 5A and B), consistent with the suppressive effects of miR-200 and with the previously ascribed roles of these genes as cytoskeletal regulators. Examination of the cells at the invading front of the 3D invasion assay revealed that miR-200 inhibited the formation of membrane protrusions (Fig 5C), which are characteristic of migratory cells and are sites of invadopodia formation (Yu & Machesky, 2012).

**MiR-200 and its targets regulate invadopodia formation**

Cell invasion is the collective outcome of a number of processes dependent on extensive cytoskeletal rearrangements and assembly of actin bundles, including the assembly of invadopodia. These are actin-rich membrane protrusions specialized for the delivery of proteases that degrade the extracellular matrix at specific sites (Sibony-Benyamini & Gil-Henn, 2012) to facilitate movement through the matrix. Given that miR-200 suppresses cell invasion (Fig 5B), we examined the potential for miR-200 to inhibit the formation of invadopodia. Invadopodia can be identified by the colocalization in dense puncta of actin with cortactin and Tks5, two proteins essential for initiation of invadopodia formation, (Courtejon et al, 2005; Bowden et al, 2006). Sites of invadopodial activity can also be identified by monitoring the local pericellular degradation of a fluorescently tagged gelatin matrix onto which the cells have been plated (Grass et al, 2012). Utilizing each of these assays, we found that miR-200 significantly reduces invadopodia formation (Fig 5D–G). We also investigated a number of miR-200 targets for their role in invadopodia formation, including WIPF1, MPRIP, and Abl2 which significantly inhibited invasion (Fig 5B), as well as cofilin2 (CFL2), moesin (MSN), and LGR4 which have all been previously implicated in promoting invasion or metastasis (Ono et al, 2008; Esteche et al, 2009; He et al, 2010; Haynes et al, 2011; Garcia et al, 2012; Gil-Henn et al, 2012; Erkutlu et al, 2013). We found that knockdown of MPRIP and WIPF1 phenocopied miR-200 in inhibiting invadopodia formation, while knocking down ABL2 reduced invadopodia numbers by about 33% (Fig 5H). CFL2 knockdown also phenocopied miR-200 (Fig 5H), consistent with a recent study reporting the role of ADF/cofilin in regulating membrane recycling essential for invadopodia formation.
miR-200 targets a cytoskeletal regulatory network

Cameron P Bracken et al

Published online: July 28, 2014
(Hagedorn et al., 2014). However, MSN1 which we had shown to mimic miR-200 in inhibiting invasion when knocked down, and to rescue the inhibition of metastasis by miR-200 when reintroduced (Li et al., 2013), had no detectable effect on invadopodia formation. Invadopodia formation was also unaffected by knockdown of LGR4 (Fig 5H).

**MiR-200 targets regulate individual steps in invadopodia formation**

Invadopodia formation proceeds in several steps, including their initiation by Src-mediated phosphorylation of cortactin and Tks5, disassembly of adjacent focal adhesions, assembly of fresh actin filaments at the site of invadopodial protrusion, and delivery of matrix metalloproteases (MMPs) to the invadopodia for membrane insertion or localized secretion (reviewed in Murphy & Courtneidge, 2011). We assessed the effect of miR-200 on each of these steps. Phosphoprotein immunoblot analysis showed that miR-200 did not affect Src activation, or the phosphorylation state of cortactin or Tks5, indicating that the effects of miR-200 on invadopodia formation were not due to effects on the initiation steps (Supplementary Fig S5). Maturation of invadopodia is accompanied by local release of MMPs, principally MMP9 and MMP14 (also known as MT1-MMP) (Poincloux et al., 2009). We assessed MMP secretion using an in-gel zymase assay, which indicated that MMP9 is strongly downregulated by miR-200, whereas MMP2, which is not associated with invadopodia function, was not detectably affected by miR-200 (Fig 6A). Furthermore, MMP9 was not indicated by Ago-HITS-CLIP to be a miR-200 target and this was verified using a luciferase reporter assay (Fig 6B), suggesting that the decrease in secreted MMP9 could be due to the loss of invadopodia formation. However, we found that miR-200 strongly repressed the level of MMP9 mRNA (Fig 6C), suggesting that miR-200 indirectly repressed MMP9 synthesis independently of its association with invadopodia. Similarly, we found that although MMP14 was not a direct target, miR-200 also repressed both MMP14 protein and mRNA levels (Fig 6D and E). The direct miR-200 targets responsible for decreasing synthesis of MMP9 and MMP14 are yet to be identified.

---

**Figure 2. Validation of miR-200 targets by luciferase reporter assay and qPCR.**

A, B Relative activities of full-length 3’UTR-luciferase reporters (A) and relative mRNA levels as measured by qPCR (B) of selected miR-200 targets in MDA-MB-231 cells transfected with control or miRNA mimics. Colored boxes below the histograms indicate the nature of the target site identified by Ago-HITS-CLIP (MM represents mismatch sites). Error bars show s.e.m. n ≥ 3; *P < 0.05; one-tailed Student’s t-test.

C, D Pooled data from the luciferase reporter assays (C) and qPCRs (D) grouped according to the nature of the target site and represented as box plots. Significance (*P < 0.05) was calculated by two-tailed t-test of mean = 1.

E Immunoblot analysis of selected targets. Relative quantitation is shown in right panel.

Source data are available online for this figure.

---

**Figure 3. Ago-HITS-CLIP reveals miR-200b targeting of EGFR via an extended 3’ UTR.**

A Histogram displaying read peaks across the EGFR 3’UTR. The y-axis shows the number of overlapping unique sequencing reads comprising the peak, and the x-axis indicates the position of the peak within the 3’UTR. The locations of potential seed sites (black arrows, 8-mers; purple arrows, 7-mers; yellow arrows, 6-mers; asterisk, central-paired) are indicated. Sequence alignment of miR-200b to the target site identified by Ago-HITS-CLIP is shown below. Black bars below the histograms indicate EGFR 3’UTRs annotated in Refseq and ESTs from GenBank.

B Measurement by qPCR of total and long UTR forms of EGFR mRNA in MDA-MB-231 cells transfected with control or miRNA mimics. n = 3; *P < 0.05; one-tailed Student’s t-test.

C Relative activity of luciferase reporters with long and short forms of the EGFR 3’UTR in MDA-MB-231 cells transfected with control or miRNA mimics. n = 3; *P < 0.05; one-tailed Student’s t-test.

Source data are available online for this figure.
miR-200 targets a cytoskeletal regulatory network

Cameron P Bracken et al

The EMBO Journal

Vol 33 | No 18 | 2014

Published online: July 28, 2014
Focal adhesions are specialized structures in which integrins provide an adhesive link between the ECM and the actin cytoskeleton (Wehrle-Haller, 2012). As the dynamic assembly and disassembly of focal adhesions are central to cell motility and disassembly of focal adhesions is an obligatory step in invadopodia formation, we also assessed the effect of miR-200 and its targets on focal adhesions. We found that the expression of miR-200 increased the number and width of focal adhesions (Fig 7) and decreased their dynamic rearrangements (Supplementary Movies S1A and S1B). Whereas knockdown of WIPF1, CFL2, and MPRIP did not affect focal adhesions, knockdown of ABL2 phencopied miR-200, increasing focal adhesion number and size (Fig 7). Together, our data indicate that miR-200 acts through overlapping subsets of targets to regulate cell invasion at multiple independent levels, including decreasing numbers of invadopodia, decreasing the synthesis of invadopodia-associated MMPs, and stabilizing cell-matrix adhesion.

**MiR-200 targets identified by HITS-CLIP reflect functionally significant targets in primary breast tumors**

Expression of the miR-200 family can become dysregulated in cancer, driving inappropriate EMT/MET processes that underlie cancer invasion and metastatic recolonization (Gregory et al, 2005a; Korpal et al, 2011; Creighton et al, 2013). To assess the potential for the involvement of the miR-200 target genes that we identified by Ago-HITS-CLIP, we assessed whether variations in miR-200 levels between cell lines and between breast cancer patients are reflected by the reciprocal variation of the network of cytoskeletal regulators that we had identified as targets of miR-200. Because CFL2, MPRIP, and WIPF1 were especially strong responders to miR-200 in luciferase reporter assays (Fig 2) and were found to promote both invadopodia formation and invasion (Fig 5), we first compared the levels of each of these to miR-200 in a panel of human breast cancer cell lines. This indicated a pronounced reciprocal relationship of each of these genes to miR-200 (Fig 8A–C), with the more epithelial cell lines expressing high levels of miR-200 and low levels of the 3 targets, whereas the more mesenchymal cell lines showed the reciprocal pattern of expression.

To more broadly assess the responses of miR-200 targets, we calculated Spearman correlation coefficients for the relationship between miR-200 and all of the targets identified by HITS-CLIP across 934 breast cancers from the TCGA dataset and 59 cell lines from the NCI-60 panel (Fig 8D and E). Because miR-200c possesses an identical seed region to miR-200b, but is expressed from a separate gene locus, we also included miR-200c in the analysis. We observed a distinct bias toward inverse correlation, supporting the notion that the cytoskeletal miR-200 targets we identify here reflect functionally relevant targets in vivo. The bias is slightly stronger in the NCI-60 dataset than in the TCGA dataset, probably because the mixture of cell types in tumor samples introduces some unrelated variation, which the cell lines of the NCI-60 panel are not affected by. In both datasets, it is evident that the tendency toward inverse correlation is strongest for targets that contain 8-mer sites in the 3′UTR and is progressively less strong for 7-mer and 6-mer sites in the 3′UTR and for sites in the coding region (CDS) (Supplementary Fig 5Aa and B). Such bias was not present when the miR-200 targets were correlated against a control miRNA (Supplementary Fig S7). This hierarchy of sites is in good agreement with our assessment of the functional effects of miR-200 on a panel of Ago-HITS-CLIP identified targets (Fig 2C and D). Since microRNAs can repress translation in addition to their effect on mRNA abundance (Carthew & Sontheimer, 2009), the inverse correlations we observe in these analyses may be even stronger at the protein level. Together, these data underscore the capacity of Ago-HITS-CLIP to identify functionally significant targets and are consistent with the notion that miRNAs regulate networks through the combination of both major effects on key genes and cumulative subtle effects on a wide range of targets.

**Discussion**

We have found hundreds of interaction sites for miR-200a and miR-200b, of which the majority were in 3′UTRs of mRNAs. Expression profile analysis, luciferase reporter assays, and correlation of expression between miR-200 and its targets in NCI-60 cell lines and the TCGA breast cancer dataset validate the capacity of Ago-HITS-CLIP to find functionally significant targets. We further confirm the hierarchical nature of target site effectiveness, with longer seed matches (8-mer and 7-mers) within 3′UTRs responsible for stronger effects of miRNAs on specific targets, while a plethora of shorter seed-matched and non-canonical sites (Chi et al, 2012; Loeb et al, 2012) provide more subtle regulatory effects on a broader range of genes. In view of the uncertainty regarding the efficacy of miRNA-binding sites in coding regions (Gu et al, 2009; Forman & Coller, 2010), it is interesting that the effect of miR-200 on CDS sites (at least for 8-mer seed matches) is sufficient to produce a bias toward negative correlations across the multiple cell lines within the NCI-60 panel, supporting the functionality of CDS sites.

Our identification of functional sites revealed a model of hierarchical network regulation that is the most extensive validation to date of the concept that individual miRNAs can broadly target whole networks. It was noteworthy, however, that although miR-200 affects both invadopodia and focal adhesions, this is through distinct subnetworks. Despite their impact on invadopodia, WIPF1...
miR-200 targets a cytoskeletal regulatory network

Cameron P Bracken et al

Published online: July 28, 2014

The EMBO Journal

Vol 33 | No 18 | 2014
and CLF2 depletion did not promote focal adhesions. Instead, focal adhesions were influenced by another miR-200 target, ABL2, which is consistent with recent evidence that Abl supports focal adhesions at the cell periphery (Peacock et al., 2010). Increased focal adhesion size and number, such as was promoted by miR-200, may inhibit cell migration by preventing adhesive release and cell translocation. Thus, one mechanism for miR-200 to inhibit cell migration may be through stabilization of focal adhesions. This yields a model where miR-200 enforces the epithelial phenotype, and prevents invasion, through distinct subnetworks that coordinately inhibit invadopodia formation and stabilize focal adhesions.

Rho signaling is an important upstream regulator of the cytoskeleton and is thought to be dysregulated in cancer through the network of GEFs and GAPs that control the Rho GTPase cycle, rather than by mutation of Rho itself (Bos et al., 2007). We found that miR-200 inhibits the Rho signaling pathway through both the targeting of the Rho activators, ARHGEF3 and NET1, and the targeting of multiple downstream elements in the Rho signaling pathway: the Rho effector ROCK2; a number of direct ROCK substrates such as the non-muscle myosins MYH9 and MYH10, which are necessary for contractility; ANLN, which mediates the interaction between myosin and RhoA; and the myosin regulators MPRIP, MYPT1, and

![Image](image-url)
MYLK, which control actomyosin contractility through regulating myosin light chain phosphorylation. ROCK also phosphorylates other miR-200-targeted genes including the ERM family members moesin and radixin, which tether actin filaments to the cell membrane; CFL2 (indirectly via LIMK), which promotes reorganization of actin filaments; and several members of the mitogen-activated protein kinase (MAPK) cascade, MAP3K7, MEKK1, and MAP2K4, which mediate JNK and p38 MAPK activation in response to growth factor signaling. Many of these aforementioned miR-200 targets have been reported as positive regulators of cell migration, invasion, and metastasis (Wang et al., 2004; Ono et al., 2008; Safina et al., 2008; Estecha et al., 2009; Su et al., 2009; Wong et al., 2009; He et al., 2010; Haynes et al., 2011; Garcia et al., 2012; Gil-Henn et al., 2012; Erkutlu et al., 2013).

In addition to directly targeting multiple pathway components, miRNAs can further coordinate networks by targeting transcription factors. We identified a number of transcription factors, some of which have been previously identified either as miR-200 targets or have been implicated in EMT, invasion, and cancer progression, including ZEB1, SUZ12, STAT5B, E2F3, TCF12, CTNNB1, and several SMADs (Gregory et al., 2008a; Iliopoulos et al., 2010; Williams et al., 2012; Xia et al., 2012; Peng et al., 2012; Chen et al., 2013; Gal et al., 2008). Targeting such transcription factors extends the network influenced by miR-200, providing the capacity to control additional processes indirectly. Interestingly, though we have not defined the precise mechanism, MMPs respond strongly but indirectly to miR-200, providing another avenue through which miR-200 controls invasion and further demonstrating how miRNAs function as master regulators, affecting multiple levels of a regulatory hierarchy.

This work provides the first demonstration of a global regulatory network directly regulated by miR-200 which strongly involves, but is not exclusively limited to, a network of signals and effectors that mediate the impact of Rho in tumor cells. This ultimately influences the epithelial–mesenchymal plasticity of cells and their ability to invade and metastasize.

### Materials and Methods

#### Cell lines and cell culture

All cell lines were cultured in Dulbecco’s Modified Eagles Medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (FBS) with the exception of MDA-MB-435 and BT-549. MDA-MB-435 were maintained in Alpha Modified Eagles Medium (aMEM; Invitrogen) supplemented with 5% FBS. BT-549 were maintained in RPMI 1640 (Invitrogen) supplemented with 10% FBS.

#### Isolation of RNA and Real-Time PCR

Total RNA was extracted using TRIzol (Invitrogen) according to the manufacturer’s instructions, and real-time PCR performed using primers as listed (Supplementary Table S2). MicroRNA PCRs were
performed using TaqMan microRNA assays (Applied Biosystems). Real-time PCR data for miRNA are expressed relative to the averaged values for GAPDH, RPL32, RPL39, and HPRT. PCR data for miRNA are normalized to U6.

**Transfection and reporter assays**

For reporter assays, cells were plated in 24-well plates and cotransfected using Lipofectamine 2000 (Invitrogen) with 200 ng firefly luciferase reporter plasmid and 5 ng pCI-neo-hRL Renilla plasmid (Pillai et al., 2005) into which 3 UTRs were cloned (see Supplementary Table S3 for cloning primers). In Fig 2A, cells were cotransfected with 20 nM scrambled, miR-200a, or miR-200b microRNA mimic (GenePharma). After 48 h of incubation, cells were lysed in passive lysis buffer (Promega) and luciferase activity measured with the Dual-Luciferase Reporter Assay System (Promega) using the TD-20/20 Luminometer (Turner Designs). All reporter assays are shown as relative luciferase activities (averaged ratios of Renilla/Firefly luciferase/C0 SEM from at least three separate experiments). In Fig 3 where plasmids were not cotransfected, RNAiMAX (Invitrogen) was used to transfet miRNA mimics. For siRNA experiments, cells were transfected with 20 nM Smart Pool siRNAs (Dharmacon) using Lipofectamine RNAiMAX and analyzed after 72 h.

**Western Blotting**

Whole-cell extracts were prepared from transfected cells by Triton X-100 lysis (50 mM Hepes, pH 7.5, 150 mM sodium chloride, 10 mM...
sodium pyrophosphate, 5 mM EDTA, 50 mM sodium fluoride, 1% Triton X-100 with protease inhibitor cocktail) and 50 µg total protein fractionated on 7.5%–15% SDS–polyacrylamide gels (dependent upon protein size). Proteins were transferred onto nitrocellulose membranes and probed with the following primary antibodies: anti-CFL2 (Abnova; 110-06848), MPRIP (Sigma; HPA022901), MSN (Cell Signaling; Q480), WIPF1 (Santa Cruz; sc25533), ROCK2 (Research Diagnostics; AF4790), CRKL (Cell Signaling; 32H4), ABL2 (Sigma; WH0000027M9), MT1-MMP (Millipore; MAB3314), α-Tubulin (Abcam; ab7291). Blots were then probed with secondary antibodies (goat anti-mouse Alexa594 and goat anti-rabbit Alexa488, Li-Cor) and visualized using the Li-Cor Odyssey.

Immunofluorescence

Coverslips were prepared using 0.2% porcine gelatin, as for fluorescent matrix degradation assay. Cells were fixed at 37°C in 4% paraformaldehyde (PFA) in cytoskeleton stabilization buffer (10 mM PIPES pH 6.8, 100 mM KCl, 300 mM sucrose, 2 mM EGTA, 2 mM MgCl2) for 15 min, permeabilized with 0.1% Triton X-100/PBS for 10 min, and blocked with 5% FBS/PBS (Cortactin ab) or 3% BSA/PBS (Tks5 ab) for 20 min. Samples were incubated with 1:1,000 Cortactin primary antibody (Upstate 05-180 mouse monoclonal) in 0.1% Triton X-100/PBS, 1:500 TKS5 primary antibody (Millipore cat# 09-268) or vinculin primary antibody (Chemicon, cat#MAB3574) in 3% BSA/PBS primary antibodies overnight at 4°C and with secondary antibodies and phalloidin for 2 h. After washing, coverslips were mounted with DAKO fluorescent mounting medium. Each experiment was repeated at least three times.

Immunoprecipitation

Seventy-two hours after transfection, cells were washed with PBS containing 100 µM Na2VO4 and lysed in lysis buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 10 mM NaF, 100 µM Na2VO4, 2 mM DTT, protease inhibitors). Protein extracts were centrifuged at 16,000 × g for 15 min at 4°C, and protein was quantitated with Bradford assay. 500 µg protein lysate was incubated mixing at 4°C for 2 h with 2 µg Cortactin (Upstate) or 4G10 phosphotyrosine (Cell Signalling) antibodies. Primary antibodies were precipitated by incubation with 50 µl Protein G Dynal beads (Invitrogen) for 2 h at 4°C. Immunoprecipitates were electrophoresed on 10% SDS–PAGE gels and immunoblotted for phosphotyrosine (4G10, Cell Signalling), Cortactin (Upstate) or Tks5 (Millipore).

Rho activation

Levels of active and total Rho were determined using the Active Rho Pull-Down and Detection kit as per manufacturer’s instruction (Thermo Scientific, cat#16116).

Argonate:miRNA immunoprecipitation

MDA-MB-231 cells were grown in 20 × 10 cm dishes and transfected with 60 nM miRNA mimic (Ambion or GenePharma) using HiPerfect transfection reagent (Qiagen). After 24 h, cells were suspended in ice-cold PBS by scraping and subjected to UV cross-linking at 254 nm (Stratalinker). Cell pellets were lysed (0.1% SDS, 0.5% deoxycholate, 0.5% NP-40 with protease inhibitors, Roche) for 10 mins on ice followed by RQ1 DNase (Promega) at 37°C for 15 min with shaking. RNAse A/T1 (Ambion) was then added for further 8 min, prior to the addition of EDTA (30 mM). Pellets were then spun (92,000 g) and the lysate subjected to immunoprecipitation for 2 h with a pan-anti-GO antibody (2A8, kind gift of Zissimos Mourelatos) conjugated to protein-A dynabeads (Invitrogen) using bridging rabbit anti-mouse IgG (Jackson Immunolabs). Pellets were then successively washed (0.1% SDS, 0.5% deoxycholate, 0.5% NP-40 in 1× PBS; 0.1% SDS, 0.5% deoxycholate, 0.5% NP-40 in 5× PBS; 50 mM Tris pH 7.5, 10 mM MgCl2, 0.5% NP-40), and on-bead phosphatase treatment performed for 30 min with antiparosal RNAse inhibitor (Ambion). The 3’ RNA linker (CAGACCGAGACGGG) was labeled with P32 using T4-PNK (NEB) and ligated on-bead for 1 h at 16°C with T4 RNA ligase (Fermentas). Beads were then washed as previous and treated with PNK to ligate the 5’ RNA linker (AGGAGACAGACGGGxxG, with ‘x’ representing different nucleotides for barcoding). Beads were resuspended in 4× LDS Novex loading buffer with 4% 2-mercaptoethanol, incubated at 70°C for 10 min and the supernatant loaded on Novex NuPAGE 4–12% Bis-Tris acrylamide gels (Bio-Rad). After running, the Ago–RNA complexes were then transferred to nitrocellulose and exposed to film at −80°C for 3 days. Complexes running at ~110 kDa (AGO + miRNA) and ~130 kDa (AGO + mRNA fragments) were then excised with a scalpel and resuspended (100 mM Tris pH 7.5, 50 mM NaCl, 10 mM EDTA, 4 mg/ml protease K) for 20 min at 37°C. The sample was incubated for an additional 20 min in the presence of 3.5 M urea and RNA isolated by a phenol–chloroform extraction. Sample was then run on a 10% denaturing (1:19) polyacrylamide gel and exposed to film with an intensifying screen at −80°C for 5 days. Thin bands corresponding to the Ago–miRNA (~110 kD) andAGO mRNA fragments (~130 kDa) were excised, crushed, and eluted at 37°C for 1 h (1 M NaOAc, pH 5.2, 1 mM EDTA). RNA was then precipitated overnight with ethanol, centrifuged, and dried. RNA was then resuspended in 8 µl H2O, primer added (TCCGCTTCGT C TCTG) and reverse transcription performed using SuperScriptIII (Invitrogen). PCR was then performed with the above primer and an additional primer (ACGGAGACAGATGCCG) for 25 cycles. PCR product was then run on a 10% native (1:29) polyacrylamide gel, stained with Sybr Gold (Qiagen) and bands excised over a UV light box. The DNA was then precipitated using isopropanol and a final 10 cycle PCR performed with the following primers: AATTGATACGGCAGCAGCAGACTATGATCTACCTAGGAG ACGATGCG, CAAGCAGAAGCGCATACGGCCTGTCGTCG TG. Reactions were then run on 2% metaphor agarose/TE gels and bands (~115 and 135 bp) excised corresponding to the linker sequence + RNA CLIP tag. Samples were finally purified using quick-spin columns (Qiagen) and subjected to Illumina GAII 35-bp read-length deep sequencing (GeneWorks).

Alignment of sequencing data

In-house scripts used for analysis may be accessed at https://bitbucket.org/sacgf/bracken_hits-clip_2013, and were written in perl and Python using Biopython (Cock et al., 2009) and HTSeq.
The EMBO Journal Vol 33 | No 18 | 2014  2053

Analysis of sequencing data

Reference gene annotation in gtf format was downloaded from UCSC (http://genome.ucsc.edu/cgi-bin/hgTables; Group: ‘Genes and Genes Prediction Tracks’; Track: ‘Refseq Genes’). To generate the reference control, reads were combined from separate HTS-CLIP experiments from untransfected and scrambled pre-miR transfected cells. Given the size of fragments isolated from the HTS-CLIP immunoprecipitation is ~40–60 nt, mapped reads were extended 40 nt 3’ from the 5’ end of the aligned read. Duplicate alignments were removed to produce a unique set of alignments, leaving 483,237 alignments for miR-200a, ~644,442 alignments for miR-200b, and ~560,900 alignments for the control samples. For analysis of the read alignment locations, regions of the genome were defined using the reference annotation and analysis was performed using Python scripts.

Peaks were defined using scripts as follows: read depth per base position per strand was calculated for each sample. The stranded read depth of each transfection sample was compared to the control sample at the same interval. Peaks were called where depth in the transfection sample was at least three reads, and the depth in the control sample was is zero or where depth was at least threefold greater in treatment than control. Annotation of peaks was performed by comparing the genomic location with annotated genes to assign each peak a gene name and genetic region. For each peak, the most likely microRNA target site was identified by searching the reference sequence for seed matches according to various targeting rules. The miR-200a-3p and miR-200b-3p seed sequences were obtained from miRBase (miR-200a-3p: CAGTGTTA, miR-200b-3p: CAGTTATTA). Sequence matching was performed using Python regular expressions. For each annotated transcript which overlaps a peak, the mRNA sequence was analyzed computationally to identify potential microRNA seed sites. Seed sites were classified from best to worst-match in the following order: 8-mer perfect match, 7-mer perfect match, 6-mer perfect match, central-paired (miRNA nt 4–15) with 2 mismatches, 8-mer with 1 mismatch, 7-mer with 1 mismatch, 7-mer with single nucleotide insertion or deletion (seed-bulge) interaction. The best seed site which occurred within 150 bp was assigned to the peak.

For analysis of reads by location in the gene, distances from the midpoint of each read to the nearest annotated stop codon and transcript end were calculated and grouped according to the length of the feature (eg mRNA or 3’UTR). Cumulative counts were used to produce a stacked bar chart. For gene graphs, arrows indicating potential miRNA targets were drawn using the seed matching method described above. Only peaks (and all unique reads at sites corresponding to locations with a peak in the control, miR-200a or miR-200b transfections) are shown.

Invasion and migration assays

For Boyden chamber invasion assays, 72 h after transfection, 2 × 105 cells were plated into each BD Biocoat Matrigel Invasion Chamber (8.0 μM PET) for 20 h. Cells were fixed with 10% buffered formalin, washed extensively with PBS, permeabilized with 0.1% Triton X-100, stained with DAPI, washed extensively with PBS, and mounted with DAKO mounting media. Images were taken of each filter, and all cells were counted. Identical protocols were used for migration assays using Transwells (Costar; 8.0 μM PET) without Matrigel. For 3D scratch wound invasion assays, 24 h after transfection, 2 × 105 cells were resuspended in 4.5 mg/ml Matrigel Basement Membrane Matrix (BD) in serum-free DMEM and plated into each well. Matrigel was set for 1 h at 37°C then covered in 500 μl complete media. When cells reached confluence, wells were scratched. The scratch was filled with 4.5 mg/ml Matrigel, set for 1 h at 37°C, and then covered in 500 μl complete media. The rate of wound closure was monitored by live-cell imaging using an IncuCyte (Essen Biosciences) for up to 3–5 days. All assays were performed in triplicates.

Invadopodia assays

Invadopodia were identified using 3 criteria; cortactin-actin and Tks5-actin colocalizations and degradation of fluorescent gelatin matrix.

Cortactin-actin and Tks5-actin colocalizations were performed by immunofluorescent staining with cortactin or Tks5 antibodies and phalloidin (Alexa fluor 633-conjugated) as described in Immunofluorescence, after which approximately 300 cells per variable were scored for the presence or absence of cortactin-actin (or Tks5-actin) colocalization. Three independent experiments were performed.

Fluorescent matrix degradation assay

Acid-washed (20% nitric acid, 30 min) coverslips were coated with 0.01% poly-L-lysine (Sigma) for 15 min and fixed in 0.5% glutaraldehyde/PBS for 10 min. Coverslips were inverted onto 80 μl droplets of warmed 1:8 0.1% Oregon Green 488 conjugate-gelatin (Invitrogen); 0.2% porcine gelatin for 10 min. Coverslips were washed in PBS then incubated with shaking for 3 min in 5 mg/ml NaBH4 in PBS. After rinsing in PBS, coverslips were incubated at 37°C in complete medium for 2 h. 1 × 105 cells were seeded on each coverslip in duplicate, incubated for 16 h, and processed for immunofluorescence. Images were taken for at least 200 cells per sample. The percentage of cells with invadopodia was calculated as the number of cells above dark holes in fluorescent matrix normalized to total cell number (counted by DAPI staining for nuclei). Each experiment was repeated at least three times.

Gelatin zymography

Conditioned media were collected from cells at 80–90% confluence (96 h after transfection). 2× Laemmlli buffer containing no reducing agent was added to samples of conditioned media and incubated at room temperature for 15 min. Samples were analyzed using 10% gelatin PAGE gels. Gels were first incubated in Renaturing Buffer (2.5% Triton X-100 in Milli Q water) for 30 min twice,
washed in Developing Buffer (50 mM Tris–HCl pH 7.4; 5 mM CaCl₂) for 30 min, and then incubated in Developing Buffer shaking 60 rpm, 37°C, 24 h. Gels were stained in 0.1% Brilliant Blue/40% methanol/10% acetic acid overnight and destained in 25% methanol/7% acetic acid. Each experiment was repeated at least three times.

Supplementary information for this article is available online: http://emboj.embopress.org

Acknowledgements
We thank Zissimos Mourelatos for the kind gift of the Argonaute antibody, Narelle Mancini for her technical assistance, and all our laboratory colleagues for their support and advice. This work was supported by Fellowships from the National Health and Medical Research Council of Australia to YA (631383) and GJG (1026191), from the National Breast Cancer Foundation and Florey Foundation to CPB, from the Cancer Council to PAG and by grants GTN1008327 and GTN1034633 from the National Health and Medical Research Council of Australia, the National Breast Cancer Foundation Australia, the Association for International Cancer Research and the Kids’ Cancer Project.

Author contributions
CPB, KBJ, ASY, YK-G, and GJG conceived the project; CBP, JAW, XL, BKD, MAA, MS, CN, and DWT conducted experiments; JML and ASY contributed reagents; AGB assisted with manuscript preparation; CPB, JAW, XL, DL, PAC, AT, KAP, and YK-G analyzed the data; and CPB, ASY, YK-G, and GJG wrote the manuscript.

Conflict of interest
The authors declare that they have no conflict of interest.

References
Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ (2008a) The miR-200 family and miR-205
regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 10: 593 – 601


miR-200 targets a cytoskeletal regulatory network

Cameron P Bracken et al


