MECP2 Is a Frequently Amplified Oncogene with a Novel Epigenetic Mechanism That Mimics the Role of Activated RAS in Malignancy

HHS Public Access

Author manuscript

Cancer Discov. Author manuscript; available in PMC 2016 July 01.

Published in final edited form as:

Cancer Discov. 2016 January ; 6(1): 45–58. doi:10.1158/2159-8290.CD-15-0341.

MECP2 Is a Frequently Amplified Oncogene with a Novel Epigenetic Mechanism that Mimics the Role of Activated RAS in Malignancy

Manish Neupane[1,2], **Allison P. Clark**[1,2], **Serena Landini**[1], **Nicolai J. Birkbak**[3], **Aron C. Eklund**[3], **Elgene Lim**[4], **Aedin C. Culhane**[5], **William T. Barry**[5], **Steven E. Schumacher**[1,6], **Rameen Beroukhim**[1,2,6,7], **Zoltan Szallasi**[3,8], **Marc Vidal**[1,9,10], **David E. Hill**[1,9,10], and **Daniel P. Silver**[1,2,7]

[1]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[2]Depatment of Medicine, Harvard Medical School, Boston, MA 02115, USA

[3]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kgs, Lyngby, Denmark

[4]The Kinghorn Cancer Centre, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia

[5]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, and Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

[6]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA

[7]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[8]Children's Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, and Harvard Medical School, Boston, MA 02115

[9]Center for Cancer Systems Biology (CCSB), Dana-Farber Cancer Institute, Boston, MA 02115, USA

[10]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA

Abstract

An unbiased genome-scale screen for unmutated genes that drive cancer growth when overexpressed identified *MECP2* as a novel oncogene. *MECP2* resides in a region of the X-chromosome that is significantly amplified across 18% of cancers, and many cancer cell lines have

[*]**Corresponding author:** Daniel P. Silver, Dana-Farber Cancer Institute, Smith Building, Room 868b, 450 Brookline Ave, Boston, MA 02215. Phone: 617-582-8485; Fax: 617-632-4381; Daniel_Silver@dfci.harvard.edu.

Disclosure of Potential Conflicts of Interest: MN, DEH, and DPS are inventors on a patent held by the Dana-Farber Cancer Institute related to this study.

Author Contributions

D.P.S. conceived of the study; M.N., D.E.H., and D.P.S. designed the experiments; M.N., A.P.C. and S.L. performed the experiments; M.N., A.P.C., and D.P.S. interpreted the experiments; D.E.H. and M.V. provided the lentiviral ORFeome expression library, N.J.B., A.C.E, and Z.S. analyzed the relationship of copy number and A.I. to expression; E.L. established and provided PDX models, A.C.C., W.T.B., S.E.S. and R.B. analyzed the significance and prevalence of MECP2 amplification; D.P.S. wrote the manuscript, and all authors edited the manuscript.

amplified, overexpressed *MECP2* and are dependent on *MECP2* expression for growth. *MECP2* copy number gain and *RAS* family member alterations are mutually exclusive in several cancer types. The *MECP2* splicing isoforms activate the major growth factor pathways targeted by activated RAS, the MAPK and PI3K pathways. *MECP2* rescued the growth of a $KRAS^{G12C}$-addicted cell line after *KRAS* down-regulation, and activated *KRAS* rescues the growth of an *MECP2*-addicted cell line after *MECP2* downregulation. MECP2 binding to the epigenetic modification 5-hydroxymethylcytosine is required for efficient transformation. These observations suggest that *MECP2* is a commonly amplified oncogene with an unusual epigenetic mode of action.

Keywords

MECP2; RAS; oncogene; gene amplification; epigenetic therapy

Introduction

Several lines of evidence suggest there are undiscovered oncogenic drivers in human cancers. For example, many recurrent amplifications found in human cancers do not contain known oncogenes, although these amplifications are clearly advantageous to tumor growth as indicated by their presence in multiple tumor types (1). Further, the recent comprehensive TCGA analysis of adenocarcinoma of the lung found significant subsets of tumors with activated MAPK or PI3K pathways without identifiable genomic alterations that explain these cancer-promoting signaling events (2), suggesting the existence of unrecognized oncogenes. In high-grade serous ovarian cancer, 45% of the tumors in the TCGA dataset have known alterations activating PI3K/RAS signaling, but it is not apparent what alterations play this or an analogous role in the remaining 55% (3).

To search for novel oncogenes that contribute to malignant transformation when overexpressed in their wild-type form, we performed an unbiased genome-scale screen. In this screen, we employed a well-studied model system in which the combination of the SV40 early region (encoding SV40 Large and Small T antigens), hTERT, and activated RAS has been shown to be sufficient to transform a number of primary human epithelial cell types (4). Many investigations have defined the role of various oncogenes and tumor suppressors by leaving out one or more of these elements, and substituting a genetic alteration present in human cancers. In our screen, rather than substituting an individual alteration for one of these elements, we added an expression library and selected cells transformed as a result.

We generated primary breast epithelial cells (HMECs) that expressed *hTERT* and the SV40 early region (“N minus RAS” cells). These cells are unable to grow in an anchorage-independent fashion in soft agar, a surrogate endpoint for transformation; the addition of activated *RAS* to these cells allows robust soft agar growth as well as their growth as tumors in xenografts in immunocompromised mice (5). For our screen, we substituted a lentivirus-based expression library (6) for activated *RAS* and identified library members that allowed these cells to grow in soft agar.

We found that one of the fifteen screen hits, *Methyl CpG Binding Protein 2* (*MECP2*), is amplified in a significant fraction of human malignancies and selected it for further study. *MECP2* is an X-linked gene, which when mutated causes the autism spectrum disorder, Rett Syndrome (7); MECP2 has no well-described role in malignancy. *MECP2* encodes a protein known to exhibit epigenetic control functions. It binds methylated CpG DNA sequences (8) and attracts the Sin3a repressor and histone deacetylase complexes (9) thus maintaining local chromatin surrounding a methylated CpG island in a less active state. Recent studies show that it also acts as a transcriptional activator (10, 11), likely through binding to another epigenetic modification of DNA, 5-hydroxymethylcytosine (12).

Epigenetic processes play important roles in human cancers (13). Tumors often exhibit global DNA hypomethylation, but at the same time reveal promoter hypermethylation (14). These alterations are thought to reflect gene expression changes important for tumorigenesis, including the repression of tumor suppressors. Recent genomic analyses of human tumors have revealed cancer-specific alterations in genes encoding epigenetic modifiers, including genes whose products function in DNA methylation, histone modifications and chromatin remodeling (15). Given the importance of epigenetic processes in human cancer, attempts have been made to alter various epigenetic processes for therapeutic purposes. In general, epigenetic therapeutics such as DNA methylation inhibitors or HDAC inhibitors, alone or in combination, are thought to have beneficial effect by re-establishing the expression of tumor suppressors repressed by hypermethylation or repressive chromatin marks. In most circumstances, the identities of the tumor suppressors that may be targeted by such therapy is not known, nor are the factors that ultimately caused their repression during tumorigenesis. The observations reported below suggest another possible mechanism of action for epigenetic therapy, which is based upon targeting of the epigenetic mode of action of a dominant oncogene.

Here, we show that the two splicing isoforms of *MECP2* activate distinct growth factor pathways in a manner that recapitulates the major oncogenic functions of activated *RAS*. The ability of MECP2 to bind the epigenetic mark 5-hydroxymethylcytosine is important for MECP2 to confer anchorage-independent growth. *MECP2*-overexpressing cell lines derived from human cancers are addicted to *MECP2* expression, suggesting *MECP2* may be useful as a therapeutic target. In sum, these experiments identify *MECP2* as a previously unrecognized oncogene with an unusual epigenetic mode of action that has potential therapeutic implications.

Results

A Genome-Scale Screen Identifies *MECP2* as a Potential Oncogene

Primary human mammary epithelial cells (HMECs) (16) were transduced with retroviruses expressing SV40 Large T, Small T, and *hTERT* (referred to here as “N minus RAS” cells). N minus RAS cells become fully transformed and capable of anchorage-independent growth with the addition of activated *RAS* (4, 5). To screen for new oncogenes that could substitute for activated *RAS*, we started with a genome-scale lentiviral expression library (6) containing over 16,000 unique open reading frames (ORFs) in an arrayed format. Twenty-one pools of approximately 800 clones each were packaged as vesicular stomatitis virus G

(VSV-G) pseudotyped lentiviruses. Cells infected with the 21 pools, GFP-infected negative control cells, and $HRASV^{12}$-infected positive control cells were plated in soft agar. After three weeks of growth, the positive control $HRASV^{12}$- infected N minus RAS HMEC plates had numerous visible soft agar colonies, the negative control GFP-infected N minus RAS HMECs had no visible colonies, and plates from 13 of the 21 pools from the ORFeome expression library also contained visible colonies. These colonies were picked, expanded, and DNA was extracted. Library members present in these colonies were identified by PCR using primers flanking the ORF, and sequencing. Despite infection at a multiplicity of infection of 0.3 as judged by drug resistance, many of these colonies had multiple proviruses, perhaps because of frequent silencing or low expression of many proviruses (17). To validate the ability of ORFs recovered in this screen to transform N minus RAS HMECs, lentiviruses corresponding to each ORF identified by PCR and sequencing were packaged individually and used to infected N minus RAS HMECs, and growth in soft agar was assessed.

In this manner, this screen identified 15 genes capable of conferring anchorage independent growth upon N minus RAS cells (Table S1). Of these 15 validated hits, 13 were present in only a single colony, 1 was identified in 2 colonies, and 1 was identified in 6 colonies; based on these data, the screen is far from achieving saturation. We scrutinized these 15 genes for evidence of recurrent overexpression or amplification in human cancers by interrogating publicly available databases. One screen hit, OTX2, has been previously identified as an oncogene amplified in childhood medulloblastoma (18–21), and thus served as a proof of concept that this screen could identify oncogenes that participate in transformation when overexpressed in their wild-type state. Another of the screen hits, *MECP2*, was selected for further study because its high rate of amplification across the TCGA collection of tumors strongly suggested relevance for human cancer. It resides in an amplicon on the long arm of the X-chromosome (Xq28) that does not contain any known oncogene.

MECP2 is Frequently Amplified and Overexpressed in Human Cancers

To investigate the prevalence of *MECP2* amplification in human cancers, we analyzed 9221 human tumor samples from the TCGA. Figure 1A shows the overall frequency of amplification of *MECP2* across a number of human cancer types in the TCGA collection determined by GISTIC (22). Q-values (false discovery rate-corrected significance of amplification frequency) below 0.25 suggest that *MECP2* is significantly amplified above the background rate and that the presence of the amplified locus is enriched by selective pressure. The q-value of the amplicon containing *MECP2* across the entire TCGA collections of cancers is 2.4×10^{-24}, demonstrating clear selective pressure favoring tumor cells containing this amplicon. The significance of amplification across Xq in the entire TCGA dataset of 9221 cancers, and heat maps of amplification for selected cancer types in the TCGA collection are shown in Figure 1B. The q-value for amplification is most significant in the chromosomal region that contains MECP2 (green dotted line). The region containing *MECP2* is the only amplicon on the X chromosome to have a significant q-value across all human cancers (Broad Institute TCGA Copy Number Portal, Table S2).

The effect of *MECP2* amplification on *MECP2* mRNA levels is assessed in Figure 1C for some cancer types with high rates of *MECP2* amplification, and in Figure S1A for other cancer types. In cancer types with significant frequencies of *MECP2* amplification, copy number gain of *MECP2* is correlated in a highly significant manner with increased *MECP2* mRNA level.

MECP2 expression is silenced on the inactive X chromosome in females (23); if the *MECP2* amplicon in tumors confers selective advantage because of higher *MECP2* expression, the *MECP2* allele on the active X chromosome would be amplified in preference to the inactive allele in tumors arising in women. We observed a pattern of amplification that confirmed preferential amplification of the active allele in triple negative breast cancer (TNBC) and ovarian cancer (Figure S1B).

Based upon the data above, TNBC, lung, and ovarian cancer were chosen for further study. Figure 1D shows western blot analysis of whole cell lysates obtained from the cell lines with and without *MECP2* amplification representing these cancer types (Table S3). We have also assayed patient derived xenografts (PDXs) (24) established from TNBCs for MECP2 overexpression by western blot; 4 of 13 (31%) consecutive TNBC PDXs overexpressed MECP2 (Figure 1D), in accord with the 33% overall frequency of *MECP2* amplification in this breast cancer subset in the TCGA collection (Figure 1A).

MECP2 Overexpression Transforms Primary Cells Containing the SV40 Early Region and *hTERT*

The *MECP2* gene expresses two splicing isoforms that differ by inclusion of the second exon, resulting in a long isoform, termed e1, that consists of 21 unique amino acids at the amino terminus followed by a 477 amino acid shared region, and a short isoform, e2, that has 9 unique amino acids at the amino terminus attached to the same 477aa shared region. In most cell types, the expression of e1 is higher than that of e2 at the protein level.

The expression library used for our screen contained only the shorter *MECP2* splicing isoform e2; infection of N minus RAS cells with a single lentivirus encoding *MECP2* e2 induced anchorage-independent growth to an extent similar to that caused by infection with activated *HRAS*, while the longer isoform, e1, was inactive in this regard (Figure 2A, left). To rule out the possibility that the observed transformation of N minus RAS HMEC cells by *MECP2* was attributable to some unique property of N minus RAS HMECs, we constructed N minus RAS cells from an entirely different human primary breast cell type, breast primary epithelial cells (BPECs) (25). *MECP2* e2 conferred anchorage-independent growth efficiently upon N minus RAS BPECs as well, and in this cell type, *MECP2* e1 was capable of conferring growth in soft agar, albeit at considerably lower efficiency than the short isoform (Figure 2A, right). Soft agar colony size in N minus RAS HMECs that express *MECP2* e2 was approximately equal to that induced by activated *HRAS* (Figure S2A), whereas in N minus RAS BPECs, colonies induced by *MECP2* e2 overexpression were larger than those induced by activated *HRAS* (Figure S2B). Lastly, in N minus RAS BPECs, colonies induced by *MECP2* e1 overexpression were much smaller than those induced by *MECP2* e2 or activated RAS in these cells (Figure S2B).

To determine whether the expression of *MECP2* was sufficient to allow growth of N minus RAS cells as tumors in nude mice, a series of N minus RAS cells infected with either of the two splicing isoforms of *MECP2*, both isoforms together, activated *RAS*, or a GFP control, were injected into the flanks of nude mice. In accord with previous results (5, 25), N minus RAS HMECs infected with a retrovirus expressing activated *RAS* were able to form tumors in nude mice, but N minus RAS HMECs infected with a virus encoding GFP were not. N minus RAS HMECs infected with both the long and short forms (e1 and e2) formed tumors in nude mice (Figure 2B, 2C), whereas N minus RAS HMECs infected with either isoform alone were unable to form tumors. To investigate the tumor-forming abilities of the *MECP2* isoforms in a different cellular context, N minus RAS BPECs were employed. N minus RAS BPECs infected with a vector expressing activated *RAS* are known to be at about four orders of magnitude more tumorigenic in nude mice than are N minus RAS HMECs expressing activated *RAS* (25). In the context of N minus RAS BPECs, each isoform of *MECP2* allowed the growth of tumors in nude mice, though N minus RAS BPECs expressing both isoforms had a higher take rate and faster tumor growth (Figure 2B, 2D).

The histology of tumors induced by *MECP2* isoforms was similar to those induced by *HRAS* (Figure S3A, B). In accord with previously published results (25), N minus RAS HMECs expressing *HRAS* gave rise to invasive ductal carcinomas with significant areas of squamous differentiation (Figure S3B), while N minus RAS BPECs gave rise to tumors that were morphologically similar, but the latter contained more limited areas of squamous differentiation (Figure S3A). N minus RAS HMECs and BPECs expressing *MECP2* displayed histologies similar to those seen with *HRAS* in the two cell types (Figure 2E, 2F); further, in the case of BPECs, there was no histologic difference seen in tumors arising after infection with each *MECP2* isoform (Figure S3A).

The Transformation Function of MECP2 is Dependent on its DNA Binding Ability

Several functional domains of MECP2 have been defined. In both MECP2 splicing isoforms, the methyl DNA-binding domain (MBD) lies near the amino-terminal end, followed by a transcription repression domain responsible for binding HDAC complexes (Figure 2G). Three separate mutations in the MBD region of the *MECP2* gene, R106W, R111G, and F155S, completely eliminate DNA-binding and cause Rett Syndrome (26–28). These mutations were tested for their transforming ability in the context of the *MECP2* e2 isoform. They all prevent the *MECP2* e2 isoform from conferring anchorage-independent growth upon N minus RAS HMECs despite expression levels similar to wt MECP2 e2, as does a truncating mutation located in the transcription repression domain (Figure 2H, I).

The *MECP2* mutation, R133C, causes a less severe form of Rett Syndrome than other mutations in the DNA binding region of MECP2 (29). R133C prevents MECP2 binding to 5-hydroxymethylcytosine (5hmC), an epigenetic mark associated with actively transcribed genes; however, it largely preserves binding to 5-methylcytosine (5mC) (12). R133C in the context of the *MECP2 e2* isoform was able to confer anchorage-independent growth upon N minus RAS HMECs at only about 10% of the efficiency of wild-type *MECP2 e2* despite expression levels equivalent to wt MECP2 e2 (Figure 2J). This observation suggests that

binding to 5hmC, and perhaps the activation of gene expression, is important for the transforming ability of MECP2.

MECP2 and activated *RAS* have similar functions in human tumors

Because *MECP2* was isolated in a screen in which it was able to substitute for the transformation function of activated *RAS*, we tested to what extent *MECP2* could substitute for activated *RAS* in other situations. Certain cultured human cancer cell lines that contain activating *KRAS* mutations are dependent upon the continued presence of *KRAS* for growth ("RAS addiction") (30, 31). We tested whether complementation with exogenous *MECP2* could rescue growth and survival defects in such a cell line after downregulation of *KRAS* expression. As shown in Figure 3A, the expression of the *MECP2* isoforms rescued growth substantially in a dose-dependent fashion after down-regulation of *KRAS* in H358, a $KRAS^{G12C}$ -addicted lung cancer cell line, using an shRNA targeting the KRAS-3'UTR. This shRNA had been previously validated as not having significant off-target effects in this cell line (designated as "K-RasC" shRNA in (30)).

If *MECP2* amplification and *RAS* activating mutation confer similar functions during tumorigenesis, there would be little or no selective advantage conferred by the presence of both alterations in a given tumor. Therefore, there may be fewer tumors that contain both alterations than would be predicted by chance, that is, there may be a mutual exclusivity relationship between the two alterations. We investigated the relationship of *MECP2* amplification to the activation of RAS family members across human tumor types and within individual tumor types. In tumor types with the highest rates of RAS mutation, pancreatic (virtually all *KRAS* mutated) and colon adenocarcinoma (roughly 50% mutated RAS), there are few tumors with amplification of *MECP2* (GISTIC (22) detects no *MECP2*-containing amplicons in those tumor types). In contrast, the cancer types with the lowest rates of RAS mutation tend to have significant numbers of cancers with *MECP2* amplification; for example, ovarian and triple negative breast cancers both have RAS mutations in less than 1% of tumors, but have high rates of *MECP2* amplification, 38% and 33% respectively (Figure 1A).

Uterine carcinoma demonstrates a different type of mutual exclusivity relationship. In the TCGA set of uterine carcinoma samples, there are high rates of *MECP2* copy number gain (15%) and high rates of *KRAS* mutation (21%), but statistically significantly fewer cases with both alterations than would be expected by chance (log odds ratio −1.215, p=0.027, Figure 3B). Uterine cancers can be subcategorized by the presence or absence of microsatellite instability, *POLE* mutation, and the level of copy number abnormalities (CNAs); in general, tumor with microsatellite instability or POLE mutation have high numbers of point mutations and few CNAs, while *POLE* wt, microsatellite stable tumors have higher CNAs (32). These subtypes of uterine cancers in the TCGA dataset were shown to vary significantly by *MECP2* status ($p<0.0001$), such that tumors with *MECP2* copy number gain or amplification were more likely to be in the CN high subtype, and less likely to be in the MSI positive subtype. Conversely, tumors with *KRAS* mutation were significantly less likely to be in the CN high subtype, and significantly more likely to be in the MSI positive or *POLE* mutated subtype ($p<0.0001$, See Supplemental Data: Analysis of

Uterine Cancer Subtypes). Therefore, the subtypes of uterine cancer drive the mutual exclusivity of *MECP2* copy number gain or amplification and *KRAS* mutation in uterine cancer.

There are also mutual exclusivity relationships between *MECP2* copy number gain and RAS alteration in some cancer types with relatively high rates of both alterations within the same subtype. Unlike uterine carcinoma, cervical cancer does not have readily identifiable subtypes (33). There is no overlap between cases with *MECP2* gain or amplification and those with *KRAS* mutation in the TCGA collection of cervical carcinoma, a statistically significant mutual exclusivity result (log odds ratio <−3, Fisher Exact Test, p=0.041) (Figure 3B). In this tumor type, the MAPK pathway is sometimes activated by amplification or mutation of *MAPK1* (ERK2) (33). These alterations of *MAPK1* are also statistically significantly mutually exclusive with *MECP2* amplification (log odds ratio <−3, p=0.041), and there is no overlap between cases with *MAPK1* amplification or mutational activation and cases of *MECP2* gain or amplification, or *KRAS* mutation (Figure 3B), strongly suggesting these alterations are functionally redundant with one another. Similarly, head and neck cancers, a group of squamous cell carcinomas, has a high percentage of cases with *MECP2* copy number gain (30%) and some cases with RAS mutation or amplification. In this cancer type, there is a statistically significant mutual exclusivity between *MECP2* copy number gain and the union of all RAS gene mutations and amplifications (log odds ratio −0.845, Fisher Exact Test, p=0.020) (Figure 3B), again suggesting *MECP2* copy number gain and RAS activity are functionally redundant.

There are other cancer types with trends toward mutual exclusivity of *RAS* mutation and *MECP2* amplification, for example adenocarcinoma of the lung and bladder carcinoma, but the relatively small percentage of one or the other alteration requires more cases be subject to genomic analysis to have enough power to determine if these relationships are statistically significant. Taken as a whole, across all human tumor types, within a cancer type with multiple subtypes, and within monomorphic tumor types with relatively high rates of both RAS activation and *MECP2* amplification, there is a tendency for tumors to have either RAS family activating mutations or *MECP2* amplification, but not both, suggesting functional redundancy.

MECP2 Isoforms Activate Growth Factor Pathways

Given that *MECP2* is able to substitute for activated *RAS* in a transformation assay as well as in a *RAS*-addicted cell line, and the mutual exclusivity analysis above between *MECP2* copy number gain and RAS alteration in several cancer types, we investigated the possibility that *MECP2* isoforms activate the major RAS-induced growth factor pathways, the MAPK pathway and PI3K pathway. To investigate the state of MAPK pathway activation, the persistence of MAPK signaling after growth factor deprivation was assessed, a standard technique employed to determine if there is cell-intrinsic activation of this pathway (34, 35). The *MECP2* short isoform e2, but not the long isoform e1, prolonged the presence of phosphorylated ERK1 and ERK2 in N minus RAS HMECs after growth factor deprivation (Figure 4A and Figure S4). This effect required an intact MECP2 DNA binding domain (Figure 4B). To determine at what level along the signaling cascade the MECP2 e2 isoform

activates the MAPK pathway, a cRAF affinity precipitation assay (36) was used to investigate whether MECP2 increases the amount of active, GTP-bound RAS. Figure 4C, upper panel, shows that MECP2 does not increase the amount of activated, GTP-bound RAS, suggesting that the MAPK pathway is activated by MECP2 at a level below RAS in the signaling pathway. Increased MEK phosphorylation was detected in *MECP2* e2-transformed cells, as was increased phosphorylation of the downstream MAPK pathway proteins, p90RSK and ELK1 (Figure 4C bottom panel). Together, these results indicate that the short isoform of MECP2 activates the MAPK signaling pathway below RAS, but at or above the level of the MEK proteins.

To assess PI3K pathway activation, cells were deprived of growth factors and persistence of PI3K signaling monitored (Figure 4D top), or deprived of growth factors for four hours, and then stimulated with recombinant EGF for 15 minutes, and re-induction monitored (figure 4D bottom), a standard assay of PI3K pathway activation (37–42)(a timecourse of EGF treatment is shown in Figure S5). PI3K signaling as assessed by p-AKT (S473) is sustained after 5 minutes of growth factor deprivation in N minus RAS cells infected with lentiviruses expressing either *MECP2* isoform or activated RAS, but not in the same cells infected with a lentivirus expressing GFP. In addition, EGF treatment of *MECP2* or RAS expressing N minus RAS HMECs compared with the same cells transduced with a vector expressing GFP shows increased levels of p-AKT and activated downstream PI3K pathway proteins p-4EBP1 and p-S6. These observations are consistent with prior reports of experiments with neuronal cells demonstrating that the MECP2 long isoform e1 stimulated the PI3K pathway (43, 44). In neurons, PI3K pathway induction by MECP2 has been shown to involve increased BDNF expression (44, 45) and/or the regulation of IGF-1 signaling (46, 47).

Human Tumor Lines with *MECP2* Overexpression Are Addicted to MECP2

Oncogene addiction, the need for continued expression of an oncogene for maintenance of the malignant phenotype (48), is necessary for therapy directed at that oncogene to be effective. The pro-tumorigenic function of *MECP2* could be necessary only during the early steps of transformation, or *MECP2* could be required on an ongoing basis to exert its pro-oncogenic function. To distinguish between these possibilities, three non-small cell lung cancer (NSCLC) cell lines and three breast cancer cell lines were selected for further study. The non-small cell lung cancer cell lines, NCI-H1755 and NCI-H23, both have copy number gain of *MECP2* and exhibited appreciable *MECP2* expression, while NCI-H1437 did not express detectable *MECP2* as assessed by western blot (Figure 1D). The two NSCLCs expressing *MECP2* demonstrated significant growth inhibition in response to infection with lentiviruses bearing either of two *MECP2*-directed shRNAs. However, the cell line that did not express *MECP2*, NCI-H1437, did not reveal significant growth inhibition (Figure 4E). Similarly, two human breast cancer cell lines that have *MECP2* copy number gain and expressed significant amounts of *MECP2*, MDAMB468 and BT549 (Figure 1D), revealed considerable growth inhibition when infected with either of two lentiviruses expressing shRNA directed at *MECP2*, while the low-expressing breast cancer cell line ZR75-1 showed little growth inhibition (Figure 4E). A number of other cell lines with high MECP2 protein expression (Figure 1D), including the breast cancer cell line BT20, and the lung cancer cell

lines NCI-H2228 and NCI-H522, also show growth inhibition when MECP2 expression is inhibited (Figure S6).

To determine if the shRNA effects seen in Figure 4E are attributable to off-target effects, the lung cancer line NCI-1755 was infected with lentiviruses encoding GFP, the cDNA of the *MECP2* short isoform e2, the cDNA of the *MECP2* long isoform e1, or viruses encoding the cDNA of both isoforms. These cell lines were then infected with a lentivirus expressing an shRNA targeting the 3' UTR of *MECP2*, which downregulates both endogenous isoforms of *MECP2* (Figure 4E) but would not be expected to downregulate expression of the cDNAs. Figure 4F shows that the long isoform of *MECP2* did not rescue the inhibition of proliferation of NCI-H1755 caused by sP2–3'UTR, the short isoform partially rescued proliferation, but together, both isoforms almost completely rescued the proliferation inhibition caused by shRNA-mediated down-regulation of both endogenous *MECP2* isoforms, showing that the proliferative defects caused by this shRNA in Figure 4E are not a result of off-target effects.

We sought to determine if activated *RAS* could substitute for *MECP2* in an *MECP2*-addicted cell line. We found that expression of activated $KRAS^{G12V}$ in the *MECP2*-addicted triple negative breast cancer cell line MDAMB468 substantially rescued the growth inhibition seen upon downregulation of *MECP2* (Figure 4G). Thus, *MECP2* can rescue the growth of cell line addicted to activated RAS after RAS downregulation (Figure 3A), and activated RAS can rescue the growth of a cell line addicted to *MECP2* after *MECP2* downregulation (Figure 4G), demonstrating the complementary relationship between RAS activity and *MECP2* function.

Discussion

MECP2 is a Frequently Amplified Oncogene with a Unique Mechanism of Action

A genome-scale screen identified *MECP2* as a gene that substitutes for activated *RAS* to allow anchorage-independent growth. *MECP2* is frequently amplified in a number of human tumor types, and many cell lines derived from human tumors have *MECP2* amplification and overexpression. Expression of *MECP2* rescues growth of a human tumor line addicted to activated *RAS* after down-regulation of *RAS* expression, and activated *RAS* rescues an *MECP2*-addicted cell line after *MECP2* downregulation. *MECP2* induces the MAPK and PI3K growth factor signaling pathways in common with activated *RAS*. MECP2 requires DNA binding for growth factor signaling and transformation, and is heavily dependent on binding to the epigenetic modification 5-hydroxycytosine for these activities. In one cellular context, the two splicing isoforms of *MECP2* have quite distinct activities despite substantial sequence identity, with the shorter e2 splicing isoform allowing anchorage-independent growth and activating both the MAPK and PI3K pathways, and the longer e1 isoform activating the PI3K pathway but not enabling anchorage-independent growth on its own. Together, the two splicing isoforms allowed efficient growth of xenografts in nude mice. These findings indicate that *MECP2* may function as an oncogene with an unusual epigenetic mechanism of action across a substantial number of human tumors.

The Amplification of Genes on the X Chromosome is a Potent Mechanism of Increasing Gene Expression

MECP2 is not an escape gene; it is silenced on the inactive X chromosome in females (23). The amplification of this class of X-linked genes affects gene expression more potently than amplification of other genes on a per copy basis. This effect results from the fact that in males, and after X-inactivation in females, there is only one active copy of this class of genes. For example, on average, one extra active copy of such an X-linked gene results in 200% of the normal level of expression, while in comparison, a biallelically expressed autosomal gene with one extra copy is expressed at 150% of normal levels. Three extra active copies of an X-linked gene would be expected to result in 4× the normal level of expression, while an autosomal gene with three extra copies has only 2.5× the normal level of expression. Thus, relatively modest amplification of *MECP2* that affect the active allele may have a relatively large effect on expression levels.

Autism and Growth Factor Pathways

Our screen linked wild-type *MECP2* overexpression to transformation. Inherited loss of function mutations in *MECP2* are known to be responsible for the autism spectrum disorder, Rett Syndrome (7). Though this relationship may seem surprising, the explanation may lie in connections between alterations in growth factor pathways and autism that are becoming increasingly apparent. A number of germline alterations that either increase or decrease growth factor pathway activity predispose to autism (reviewed in (49)). For example, there is a significant overrepresentation of autism among patients with Beckwith-Wiedemann syndrome (in which patients overexpress IGF2 or cyclin-dependent kinase inhibitor 1C) (50), in Neurofibromatosis (NF1) (51, 52), in the PTEN hamartoma tumor syndromes Cowden Syndrome and Bannayan-Riley-Ruvalcaba (53–57), in tuberous sclerosis (58), and SNPs for three PI3K signaling pathway genes, INPP1, PIK3CG, and TSC2, are in linkage disequilibrium with autism (59). A c-MET promoter SNP that decreases cMET expression is associated with two-fold increase in autism risk (60). Mutations in the RAS pathway scaffolding protein, CNKSR2, are associated with an autism-like disorder (61). Finally, treatment with recombinant IGF1, which crosses the blood-brain barrier and activates growth factor signaling, has been shown to have therapeutic efficacy in a mouse model of Rett Syndrome (62) and in a phase 1 clinical trial treating girls with *MECP2* mutations (63).

It is not clear how defects in *MECP2* lead to Rett Syndrome. Given the connection between autism spectrum disorders and growth factor pathway alterations, and our demonstration that the two splicing isoforms of *MECP2* each contribute to the activation of growth factor pathways, the work described here may indirectly lead to further hypotheses regarding the role of abnormal growth factor pathway activation in patients with autism spectrum disorders. Further, it is tempting to speculate that the *MECP2* isoform-specific differences in growth factor pathway induction demonstrated here may have arisen to titrate and balance the activity of different growth factor pathways by the process of alternative splicing in neurons. Because the growth factor pathways induced by MECP2 activity, the MAPK and PI3K pathways, are key pathways in human tumors, it seems likely that cells overexpressing *MECP2* by amplification or other means would have a selective advantage during tumorigenesis.

Epigenetic and Other Therapies

Experiments presented here demonstrate that several human tumor cell lines with overexpressed *MECP2* show significant growth impairment after *MECP2* down-regulation by shRNA. These observations suggest that patients with tumors overexpressing *MECP2* might benefit from therapy targeting MECP2 function. Further, the necessity of DNA binding activity of MECP2 for transformation and growth factor pathway induction suggests that interfering with the ability of MECP2 to bind DNA might have therapeutic effect. Because MECP2 binds specifically to methylated or hydroxymethylated cytosine, blocking the formation of these modified cytosines might specifically inhibit *MECP2*-transformed cells. The FDA-approved drugs 5-azacytidine and decitabine inhibit DNA methylation and therefore inhibit the formation of 5-methylcytosine and 5-hydroxymethylcytosine, and so may be therapeutic in tumors with overexpressed *MECP2*. Experiments are ongoing to test this hypothesis.

Data presented here implicate the binding of MECP2 to 5hmC as important for the cancer-related functions of MECP2. 5hmC is formed from 5mC by the action of the TET enzymes (64, 65). *TET2* is a tumor suppressor in some settings; *TET2* inactivating mutations occur commonly in myelodysplasia and myeloid malignancies. Further, neomorphic oncogenic mutations have been found in the enzymes IDH1 or IDH2 in glioblastoma, AML, chondrosarcoma, and cholangiocarcinoma that result in the accumulation of (R)-2-hydroxyglutarate (66), an abnormal metabolite that inhibits the TET enzymes, as well as other 2-oxoglutarate-dependent dioxygenases. However, TET1 also possesses oncogenic activity (67). Given the dependence of the *MECP2* transformation activity on 5hmC, it is possible that inhibition of the TET enzymes may inhibit the growth of *MECP2*-related tumors; this inhibition may be accomplished by compounds similar to (R)-2-hydroxyglutarate (68). Further insight into the mechanisms MECP2 uses to drive transformation may reveal additional therapeutic targets.

MECP2- A Novel Oncogene

MECP2 has several properties that make it an unusual oncogene. First, its two splicing isoforms cooperate in tumor formation, with MAPK pathway induction contributed uniquely by the short isoform. Second, its mechanism of action requires the epigenetic modification of cytosine, including the recently discovered modification, 5-hydroxymethylcytosine. This property may provide unusual therapeutic opportunities. Lastly, *MECP2* is a member of a small class of genes that are involved in two entirely different human diseases. When mutated, *MECP2* causes Rett Syndrome, and when amplified/overexpressed, it is involved in cancer.

Methods

Vectors

Details of lentiviral and retroviral expression vectors, shRNA vectors (with target sequences), construction of mutants (including primer sequences) and cloning are provided in Supplemental Methods section.

Virus production and transduction

Virus particles were produced by transient co-transfection of 293T cells using 3 plasmid system (for retrovirus) or 5 plasmid system (for lentivirus) as detailed in supplementary section. Further information on transduction of target cells with cDNA viruses for overexpression studies and shRNA viruses for knock-down studies is described in related subsections of Supplemental Methods.

Cell culture

The HMECs were purchased from Lonza, and the BPECs were gift from Dr. Tan Ince (University of Miami) in January 2010; no authentication of these cells has been done by the authors. The BT549, MDAMB468, ZR75-1, BT20, HCC38, MCF7, MDAMB231, MDAMB453, and T47D cell lines were purchased from ATCC, and no authentication has been done by the authors. The NCI-H1755, NCI-H23, NCI-H1437, NCI-H1395, and NCI-H2347 cell lines from the Minna and Gazdar laboratories (The University of Texas Southwestern Medical Center) were provided to the Belfer Institute/Dana-Farber through the Meyerson Lab at the Broad Institute (with permission from the originators), and were obtained by authors in November 2012. The NCI-H2228, NCI-H522, NCI-H1563, NCI-H2170, and NCI-H358 cell lines were purchased from ATCC and, NCI-H2009 and NCI-H1993 cell lines were obtained from Belfer Institute in November 2012. All of these lung cancer cell lines have been short-tandem repeat (STR)-profiled and mycoplasma tested at the Belfer Institute. The EFO21, IGROV1, SKOV3, CAOV3, OVCAR3, OVCAR4, OVCAR5, and OVCAR8 cell lines were gift from Dr. Ronny Drapkin (University of Pennsylvania) in March 2013 and have been STR-fingerprinted at Drapkin laboratory. All the cell lines employed for the experiments described in this manuscript were used from early passages (for less than 6 months) of cell stocks frozen at the time of receipt. All cell lines were maintained at 37°C with 10% CO_2. Details about culture conditions, media composition and generation and maintenance of stable cell lines are described in Supplemental Methods section.

Soft agar growth and Tumorigenicity assay

Anchorage-independent growth of cells in soft agar was determined by plating 3×10^4 cells (HMECs or BPECs derivatives) in 0.3% Noble agar (Difco). Colonies were counted after 3 weeks of growth. Xenograft study was carried out in NCr nude mice (Taconic), and the study protocol was approved by the Dana-Farber IACUC. For further details on both of these assays, see Supplemental Methods.

Cell proliferation assays

Cell proliferation was quantified by either measuring cellular ATP content by CellTiter-Glo Luminescent Cell Viability Assay (Promega), or by extracting the cell-associated crystal violet dye with 10% acetic acid.

Western blot

Cells were lysed in RIPA buffer and western blot was carried out using standard protocol. For assays involving MAPK pathway activity, HMEC derivatives were washed three times

with PBS, reset in growth factor-deprived medium (MEBM) at 37°C for 5 minutes or longer, and then lysed (see Supplemental Methods for details). For assays involving PI3K pathway activity, HMEC derivatives were washed three times with PBS, starved in MEBM for five minutes and analyzed, or starved for 4 hours, treated with EGF (2.5 ng/ml; Sigma) at 37°C for 15 minutes, and then lysed. Primary antibodies used for immunoblotting are listed in Supplemental Methods.

Bioinformatic Analysis

SNP6 array data generated by TCGA from 9221 cancers across 29 cancer types were analyzed for significantly recurrent amplifications using GISTIC 2.0 (69) and according to the methods described in (70). Further details about the samples, analysis parameters, and detailed results can be found in Broad Institute TCGA Copy Number Portal, under the analysis version "2014-07-08 stddata_2014_06_14." For details of the expression vs. copy number and A.I. analyses, please see Supplemental Methods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank A. Richardson for evaluating the xenograft pathology specimens, A. MacWilliams and T. Hao for help with the lentivirus orfeome library, T. Joshi for bioinformatics analysis, D. Livingston and members of the Livingston lab for discussions and advice. We thank all of the investigators, institutions, and patients who contributed to the Cancer Genome Atlas (TCGA) research network; the results published here are in part based upon data generated by the TCGA Research Network.

Financial Support: This work was supported by a grant from the Cogan Family Foundation (D.P.S.), DFCI startup funds from philanthropic sources (D.P.S.), the DF/HCC SPORE in breast cancer NIH P50 CA168504-01A1, a U01HG001715 (M.V. and D.E.H.), DFCI Institute Sponsored Research (M.V. and D.E.H.), and the Ellison Foundation (Boston, MA) (M.V. and D.E.H.).

References

1. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010; 463:899–905. [PubMed: 20164920]
2. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511:543–550. [PubMed: 25079552]
3. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011; 474:609–615. [PubMed: 21720365]
4. Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. Nature. 1999; 400:464–468. [PubMed: 10440377]
5. Elenbaas B, Spirio L, Koerner F, Fleming MD, Zimonjic DB, Donaher JL, et al. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes Dev. 2001; 15:50–65. [PubMed: 11156605]
6. Yang X, Boehm JS, Yang X, Salehi-Ashtiani K, Hao T, Shen Y, et al. A public genome-scale lentiviral expression library of human ORFs. Nat Methods. 2011; 8:659–661. [PubMed: 21706014]
7. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet. 1999; 23:185–188. [PubMed: 10508514]

8. Meehan RR, Lewis JD, McKay S, Kleiner EL, Bird AP. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell. 1989; 58:499–507. [PubMed: 2758464]
9. Nan XX, Ng HHH, Johnson CAC, Laherty CDC, Turner BMB, Eisenman RNR, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998; 393:386–389. [PubMed: 9620804]
10. Jordan C, Li HH, Kwan HC, Francke U. Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. BMC medical genetics. 2007; 8:36. [PubMed: 17584923]
11. Chahrour MM, Jung SYS, Shaw CC, Zhou XX, Wong STCS, Qin JJ, et al. MeCP2, a key contributor to neurological disease, activates and represses transcription. Science (New York, NY). 2008; 320:1224–1229.
12. Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell. 2012; 151:1417–1430. [PubMed: 23260135]
13. Baylin SB. DNA methylation and gene silencing in cancer. Nature clinical practice Oncology. 2005; 2(Suppl 1):S4–S11.
14. Weisenberger DJ. Characterizing DNA methylation alterations from The Cancer Genome Atlas. J Clin Invest. 2014; 124:17–23. [PubMed: 24382385]
15. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell. 2012; 22:9–20. [PubMed: 22789535]
16. Stampfer MR. Isolation and growth of human mammary epithelial cells. Methods in Cell Science. 1985; 9:107–115.
17. Sastry L, Johnson T, Hobson MJ, Smucker B, Cornetta K. Titering lentiviral vectors: comparison of DNA, RNA and marker expression methods. 2002; 9:1155–1162.
18. Boon K, Eberhart CG, Riggins GJ. Genomic amplification of orthodenticle homologue 2 in medulloblastomas. Cancer Research. 2005; 65:703–707. [PubMed: 15705863]
19. Di C, Liao S, Adamson DC, Parrett TJ, Broderick DK, Shi Q, et al. Identification of OTX2 as a medulloblastoma oncogene whose product can be targeted by all-trans retinoic acid. Cancer Res. 2005; 65:919–924. [PubMed: 15705891]
20. Yokota N, Mainprize TG, Taylor MD, Kohata T, Loreto M, Ueda S, et al. Identification of differentially expressed and developmentally regulated genes in medulloblastoma using suppression subtraction hybridization. Oncogene. 2004; 23:3444–3453. [PubMed: 15064731]
21. Michiels EM, Oussoren E, Van Groenigen M, Pauws E, Bossuyt PM, Voute PA, et al. Genes differentially expressed in medulloblastoma and fetal brain. Physiol Genomics. 1999; 1:83–91. [PubMed: 11015565]
22. Beroukhim R, Getz G, Nghiemphu L, Barretina J, Hsueh T, Linhart D, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc Natl Acad Sci U S A. 2007; 104:20007–20012. [PubMed: 18077431]
23. Hoffbuhr KC, Moses LM, Jerdonek MA, Naidu S, Hoffman EP. Associations between MeCP2 mutations, X-chromosome inactivation, and phenotype. Mental retardation and developmental disabilities research reviews. 2002; 8:99–105. [PubMed: 12112735]
24. Derose YS, Wang G, Lin YC, Bernard PS, Buys SS, Ebbert MT, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med. 2011; 17:1514–1520. [PubMed: 22019887]
25. Ince TA, Richardson AL, Bell GW, Saitoh M, Godar S, Karnoub AE, et al. Transformation of different human breast epithelial cell types leads to distinct tumor phenotypes. Cancer Cell. 2007; 12:160–170. [PubMed: 17692807]
26. Amir RE, Van den Veyver IB, Schultz R, Malicki DM, Tran CQ, Dahle EJ, et al. Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. Annals of neurology. 2000; 47:670–679. [PubMed: 10805343]
27. Yusufzai TM, Wolffe AP. Functional consequences of Rett syndrome mutations on human MeCP2. Nucleic Acids Res. 2000; 28:4172–4179. [PubMed: 11058114]

28. Free A, Wakefield RI, Smith BO, Dryden DT, Barlow PN, Bird AP. DNA recognition by the methyl-CpG binding domain of MeCP2. J Biol Chem. 2001; 276:3353–3360. [PubMed: 11035019]
29. Bebbington A, Anderson A, Ravine D, Fyfe S, Pineda M, de Klerk N, et al. Investigating genotype-phenotype relationships in Rett syndrome using an international data set. Neurology. 2008; 70:868–875. [PubMed: 18332345]
30. Singh A, Greninger P, Rhodes D, Koopman L, Violette S, Bardeesy N, et al. A gene expression signature associated with "K-Ras addiction" reveals regulators of EMT and tumor cell survival. Cancer Cell. 2009; 15:489–500. [PubMed: 19477428]
31. Scholl C, Frohling S, Dunn IF, Schinzel AC, Barbie DA, Kim SY, et al. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. Cell. 2009; 137:821–834. [PubMed: 19490892]
32. Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, et al. Cancer Genome Atlas Research N. Integrated genomic characterization of endometrial carcinoma. Nature. 2013; 497:67–73. [PubMed: 23636398]
33. Ojesina AI, Lichtenstein L, Freeman SS, Pedamallu CS, Imaz-Rosshandler I, Pugh TJ, et al. Landscape of genomic alterations in cervical carcinomas. Nature. 2014; 506:371–375. [PubMed: 24390348]
34. Zhu J, Blenis J, Yuan J. Activation of PI3K/Akt and MAPK pathways regulates Myc-mediated transcription by phosphorylating and promoting the degradation of Mad1. Proc Natl Acad Sci U S A. 2008; 105:6584–6589. [PubMed: 18451027]
35. Spangle JM, Munger K. The HPV16 E6 oncoprotein causes prolonged receptor protein tyrosine kinase signaling and enhances internalization of phosphorylated receptor species. PLoS pathogens. 2013; 9:e1003237. [PubMed: 23516367]
36. Taylor SJ, Shalloway D. Cell cycle-dependent activation of Ras. Curr Biol. 1996; 6:1621–1627. [PubMed: 8994826]
37. Zhao JJ, Gjoerup OV, Subramanian RR, Cheng Y, Chen W, Roberts TM, et al. Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase. Cancer Cell. 2003; 3:483–495. [PubMed: 12781366]
38. Zhao JJ, Cheng H, Jia S, Wang L, Gjoerup OV, Mikami A, et al. The p110alpha isoform of PI3K is essential for proper growth factor signaling and oncogenic transformation. Proc Natl Acad Sci U S A. 2006; 103:16296–16300. [PubMed: 17060635]
39. Cipriano R, Graham J, Miskimen KL, Bryson BL, Bruntz RC, Scott SA, et al. FAM83B mediates EGFR- and RAS-driven oncogenic transformation. J Clin Invest. 2012; 122:3197–3210. [PubMed: 22886302]
40. Westbrook TF, Martin ES, Schlabach MR, Leng Y, Liang AC, Feng B, et al. A genetic screen for candidate tumor suppressors identifies REST. Cell. 2005; 121:837–848. [PubMed: 15960972]
41. Niepel M, Hafner M, Pace EA, Chung M, Chai DH, Zhou L, et al. Analysis of growth factor signaling in genetically diverse breast cancer lines. BMC biology. 2014; 12:20. [PubMed: 24655548]
42. Wheeler DB, Zoncu R, Root DE, Sabatini DM, Sawyers CL. Identification of an oncogenic RAB protein. Science. 2015; 350:211–217. [PubMed: 26338797]
43. Li Y, Wang H, Muffat J, Cheng AW, Orlando DA, Loven J, et al. Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. Cell Stem Cell. 2013; 13:446–458. [PubMed: 24094325]
44. Ricciardi S, Boggio EM, Grosso S, Lonetti G, Forlani G, Stefanelli G, et al. Reduced AKT/mTOR signaling and protein synthesis dysregulation in a Rett syndrome animal model. Hum Mol Genet. 2011; 20:1182–1196. [PubMed: 21212100]
45. Chang Q, Khare G, Dani V, Nelson S, Jaenisch R. The disease progression of Mecp2 mutant mice is affected by the level of BDNF expression. Neuron. 2006; 49:341–348. [PubMed: 16446138]
46. Itoh M, Ide S, Takashima S, Kudo S, Nomura Y, Segawa M, et al. Methyl CpG-binding protein 2 (a mutation of which causes Rett syndrome) directly regulates insulin-like growth factor binding protein 3 in mouse and human brains. Journal of neuropathology and experimental neurology. 2007; 66:117–123. [PubMed: 17278996]

47. Tropea D, Giacometti E, Wilson NR, Beard C, McCurry C, Fu DD, et al. Partial reversal of Rett Syndrome-like symptoms in MeCP2 mutant mice. Proc Natl Acad Sci U S A. 2009; 106:2029–2034. [PubMed: 19208815]
48. Weinstein IB. Cancer. Addiction to oncogenes--the Achilles heal of cancer. Science. 2002; 297:63–64. [PubMed: 12098689]
49. Crespi B. Autism and cancer risk. Autism research : official journal of the International Society for Autism Research. 2011; 4:302–310. [PubMed: 21823244]
50. Kent L, Bowdin S, Kirby GA, Cooper WN, Maher ER. Beckwith Weidemann syndrome: a behavioral phenotype-genotype study. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2008; 147B:1295–1297.
51. Marui T, Hashimoto O, Nanba E, Kato C, Tochigi M, Umekage T, et al. Association between the neurofibromatosis-1 (NF1) locus and autism in the Japanese population. American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics. 2004; 131B:43–47.
52. Gillberg C, Forsell C. Childhood psychosis and neurofibromatosis--more than a coincidence? Journal of autism and developmental disorders. 1984; 14:1–8. [PubMed: 6423617]
53. Butler MG, Dasouki MJ, Zhou XP, Talebizadeh Z, Brown M, Takahashi TN, et al. Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. J Med Genet. 2005; 42:318–321. [PubMed: 15805158]
54. Greer JM, Wynshaw-Boris A. Pten and the brain: sizing up social interaction. Neuron. 2006; 50:343–345. [PubMed: 16675386]
55. Herman GE, Butter E, Enrile B, Pastore M, Prior TW, Sommer A. Increasing knowledge of PTEN germline mutations: Two additional patients with autism and macrocephaly. American journal of medical genetics Part A. 2007; 143:589–593. [PubMed: 17286265]
56. McBride KL, Varga EA, Pastore MT, Prior TW, Manickam K, Atkin JF, et al. Confirmation study of PTEN mutations among individuals with autism or developmental delays/mental retardation and macrocephaly. Autism research : official journal of the International Society for Autism Research. 2010; 3:137–141. [PubMed: 20533527]
57. Varga EA, Pastore M, Prior T, Herman GE, McBride KL. The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. Genet Med. 2009; 11:111–117. [PubMed: 19265751]
58. Harrison JE, Bolton PF. Annotation: tuberous sclerosis. Journal of child psychology and psychiatry, and allied disciplines. 1997; 38:603–614.
59. Serajee FJ, Nabi R, Zhong H, Mahbubul Huq AH. Association of INPP1, PIK3CG, and TSC2 gene variants with autistic disorder: implications for phosphatidylinositol signalling in autism. J Med Genet. 2003; 40:e119. [PubMed: 14627686]
60. Campbell DB, Sutcliffe JS, Ebert PJ, Militerni R, Bravaccio C, Trillo S, et al. A genetic variant that disrupts MET transcription is associated with autism. Proc Natl Acad Sci U S A. 2006; 103:16834–16839. [PubMed: 17053076]
61. Houge G, Rasmussen IH, Hovland R. Loss-of-Function CNKSR2 Mutation Is a Likely Cause of Non-Syndromic X-Linked Intellectual Disability. Molecular syndromology. 2012; 2:60–63. [PubMed: 22511892]
62. Castro J, Garcia RI, Kwok S, Banerjee A, Petravicz J, Woodson J, et al. Functional recovery with recombinant human IGF1 treatment in a mouse model of Rett Syndrome. Proc Natl Acad Sci U S A. 2014; 111:9941–9946. [PubMed: 24958891]
63. Khwaja OS, Ho E, Barnes KV, O'Leary HM, Pereira LM, Finkelstein Y, et al. Safety, pharmacokinetics, and preliminary assessment of efficacy of mecasermin (recombinant human IGF-1) for the treatment of Rett syndrome. Proc Natl Acad Sci U S A. 2014; 111:4596–4601. [PubMed: 24623853]
64. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324:930–935. [PubMed: 19372391]

65. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature. 2010; 466:1129–1133. [PubMed: 20639862]

66. Losman JA, Kaelin WG Jr. What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. Genes Dev. 2013; 27:836–852. [PubMed: 23630074]

67. Huang H, Jiang X, Li Z, Li Y, Song CX, He C, et al. TET1 plays an essential oncogenic role in MLL-rearranged leukemia. Proc Natl Acad Sci U S A. 2013; 110:11994–11999. [PubMed: 23818607]

68. Losman JA, Looper RE, Koivunen P, Lee S, Schneider RK, McMahon C, et al. (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. Science. 2013; 339:1621–1625. [PubMed: 23393090]

69. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011; 12:R41. [PubMed: 21527027]

70. Zack TI, Schumacher SE, Carter SL, Cherniack AD, Saksena G, Tabak B, et al. Pan-cancer patterns of somatic copy number alteration. Nat Genet. 2013; 45:1134–1140. [PubMed: 24071852]

A

Cancer Subset	Q-value	Overall Frequency of Amplification
All Cancers	$2.4x10^{-24}$	18%
All Epithelial	$1.3x10^{-22}$	20%
Ovarian	$7.3x10^{-10}$	38%
Triple Negative Breast	0.175	33%
Lung Adeno	$1.5x10^{-6}$	29%
Head and Neck	0.01	28%
Lung Squamous	$9.9x10^{-2}$	25%
Cervical	$8.8x10^{-4}$	25%
Liver	0.085	24%
Stomach	0.06	19%
Uterine	$9.3x10^{-3}$	17%

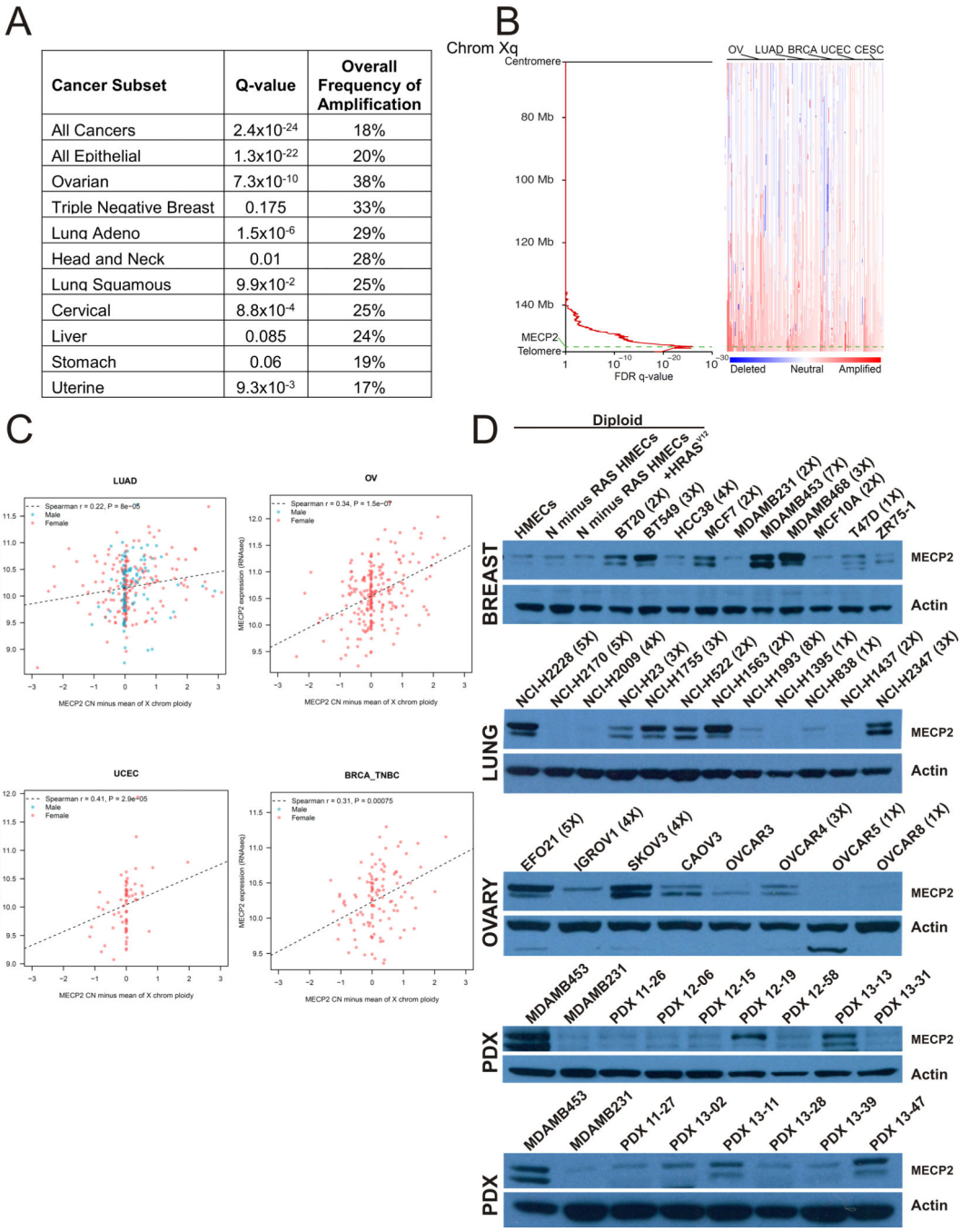


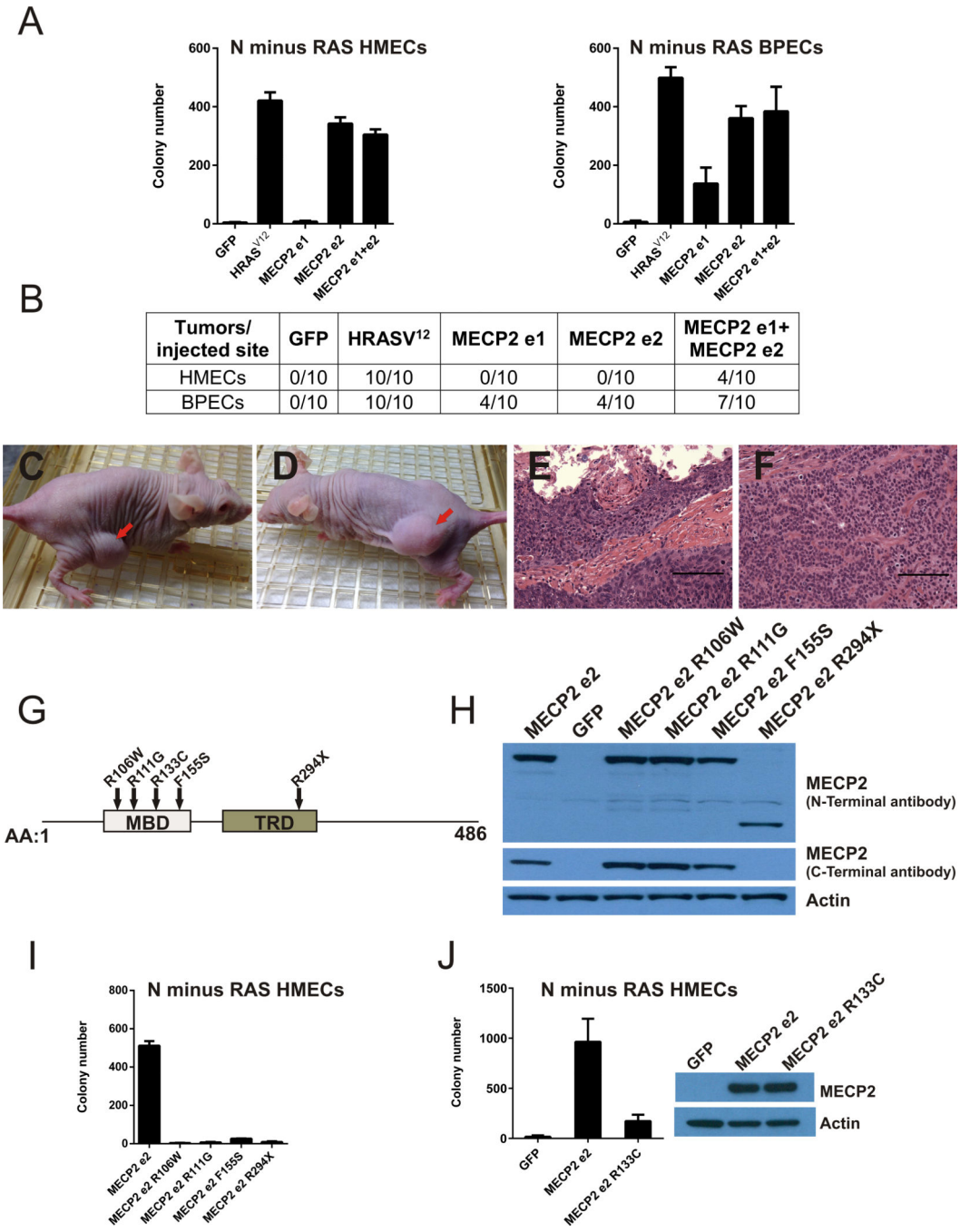
Figure 1. *MECP2* is Amplified and Overexpressed in Several Human Cancer Types

(A) *MECP2* amplification frequencies in selected cancer types are shown. Source: Broad Institute TCGA Copy Number Portal, 9221 Tumor Samples, 2014-07-08 data release.

(B) The significance of *MECP2* amplifications determined by GISTIC (Genomic Identification Of Significant Targets In Cancer)(22) across 9221 cancers in the TCGA collection is shown alongside a heat map of copy number alterations of Xq in selected human cancers with *MECP2* amplification. The dotted green line indicates the location of the MECP2 gene.

(C) The relationship of *MECP2* expression to *MECP2* copy number is shown for lung adenocarcinoma, ovarian cancer, uterine cancer, and triple negative breast cancer (TNBC) (TCGA dataset). The X axis shows the copy number in a 2mb region surrounding *MECP2* minus the average copy number of the total X chromosome, the Y axis shows *MECP2* RNA expression quantitated by RNAseq.

(D) Western blots of MECP2 protein expression in total cell lysates from human cell lines and Patient Derived Xenografts (PDXs) derived from TNBC. Diploid cell lines without *MECP2* amplification are for comparison (top); all other samples are from human cancer cell lines and TNBC PDXs. Where known, the *MECP2* copy number is in parentheses.



Tumors/ injected site	GFP	HRASV12	MECP2 e1	MECP2 e2	MECP2 e1+ MECP2 e2
HMECs	0/10	10/10	0/10	0/10	4/10
BPECs	0/10	10/10	4/10	4/10	7/10

Figure 2. Oncogenic Activities of *MECP2* Splicing Isoforms are Distinct and Require DNA-binding

(A) The ability of *MECP2* splicing isoforms to confer anchorage-independent growth to N minus RAS HMECs and N minus RAS BPECs (a different strain of primary human breast epithelial cells) in comparison to isogenic non-transformed GFP-infected cells and isogenic HRASV^{12}-transformed cells is shown. The average number of soft agar colonies from 3 replicates, +/− sd is indicated.

(B–F) N minus RAS HMECs or N minus RAS BPECs infected with viruses expressing the indicated genes were injected into the flanks of nude mice. The number of tumors observed

is shown. Ten injections were performed for each cell type (B). Representative photographs of tumor-bearing nude mice injected with N minus RAS HMEC cells infected with both isoforms of *MECP2* (C) or N minus RAS BPEC cells infected with both isoforms of *MECP2* (D) and 40× photomicrographs of tumor H&E histology are shown from N minus RAS HMECs expressing both isoforms of *MECP2* (E) or N minus RAS BPECs expressing both isoforms of *MECP2* (F). Scale bars, 0.1 mm.

(G) The location of structural features and mutations in the context of the *MECP2* short isoform e2.

(H) Protein expression of mutant and wt alleles in the context of *MECP2* e2 from whole cell lysates of N minus RAS HMECs.

(I) Soft agar growth of N minus RAS HMECs expressing *MECP2* e2 mutations that completely abrogate the ability of MECP2 to bind DNA.

(J) Soft agar colony counts for N minus RAS HMECs expressing GFP, wt *MECP2* e2, or *MECP2* e2 bearing the mutation R133C, which prevents MECP2 binding to 5hmC and leaves binding to 5mC largely intact. The average number of soft agar colonies from 3 replicates, +/− sd are indicated. Inset: MECP2 immunoblot of whole cell lysates prepared from cells used in soft agar assay.

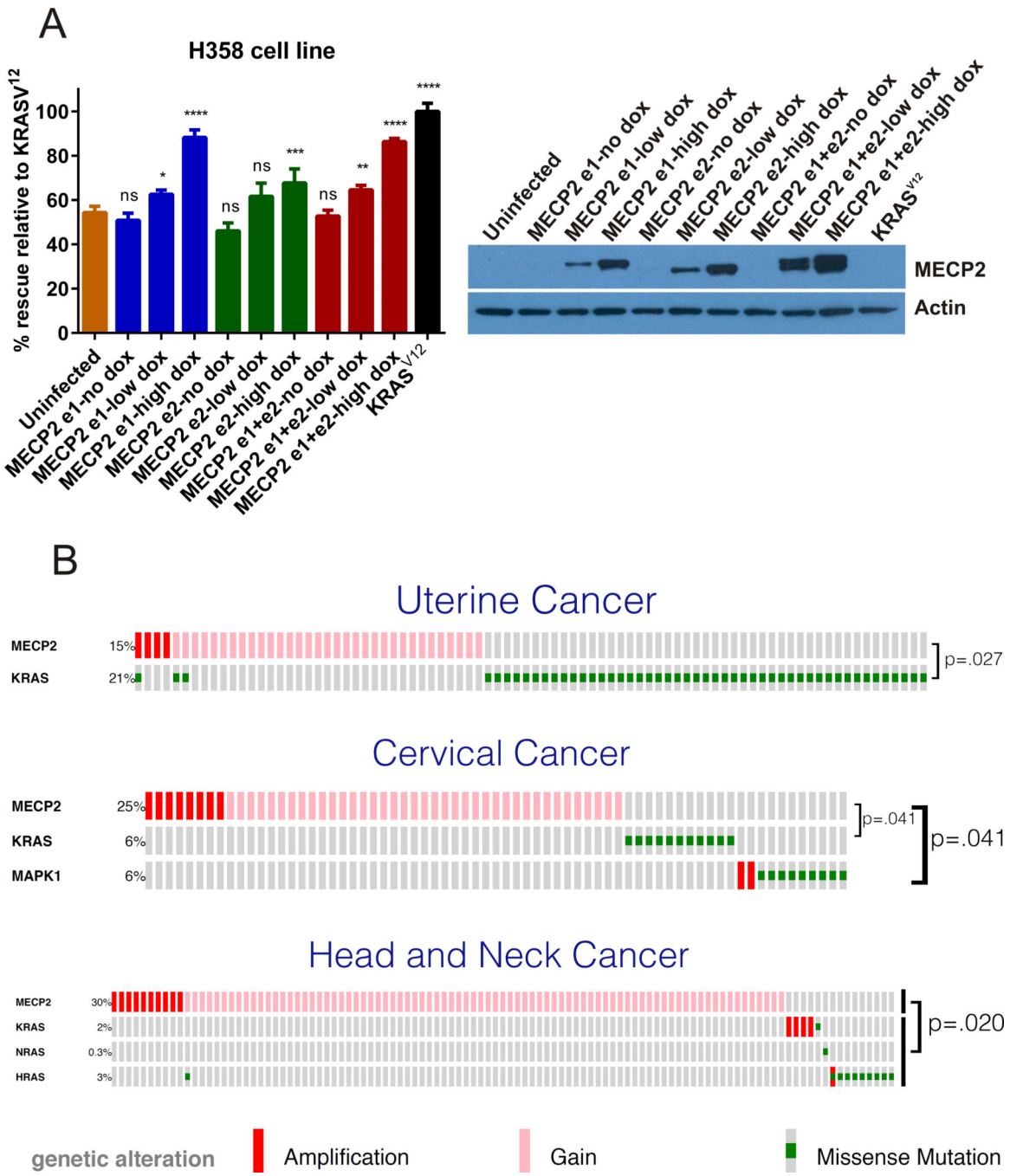


Figure 3. The Functions of *MECP2* and Activated *RAS* are Similar in Human Cancers

(A) The *KRAS*-addicted cell line H358, a cell line derived from a human adenocarcinoma of the lung, was infected with doxycycline-inducible constructs expressing the *MECP2* e1 isoform, the e2 isoform, both isoforms, or with a virus expressing activated *KRAS*. Increasing levels of *MECP2* expression were induced by the use of higher concentrations of doxycycline (shown in MECP2 immunoblot on the right). All cell lines were then infected with a virus expressing a validated shRNA targeting *KRAS* (sh*KRAS*#3'UTR), and 12 days later, cell numbers were measured by quantitating ATP with the CellTiter-Glo Luminescent

Cell Viability Assay. Mean ±SD data from 4 independent viral infections (n=4) were analyzed using one-way ANOVA followed by post hoc Dunnett's multiple comparison test. Asterisks indicate statistically significant difference between each experimental group and uninfected control (****: p<0.0001, ***: p=0.0001, **: p=0.0030, *: p=0.0237, ns: non-significant, p>0.05)

(B) Graphical summary (Oncoprint) demonstrating mutual exclusivity of *MECP2* amplification or copy number gain, RAS family alterations, and *MAPK1* (ERK2) alterations in uterine, cervical, and head and neck cancer. The TCGA dataset (242 cases of uterine cancer, 191 cases of cervical cancer, and 302 cases of head and neck cancer with both sequencing and copy number data) were analyzed for mutual exclusivity using the cBio Portal for Cancer Genomics; p values for mutual exclusivity are indicated (Fisher Exact Test). Only tumors with alterations in the indicated genes are shown in the figure.

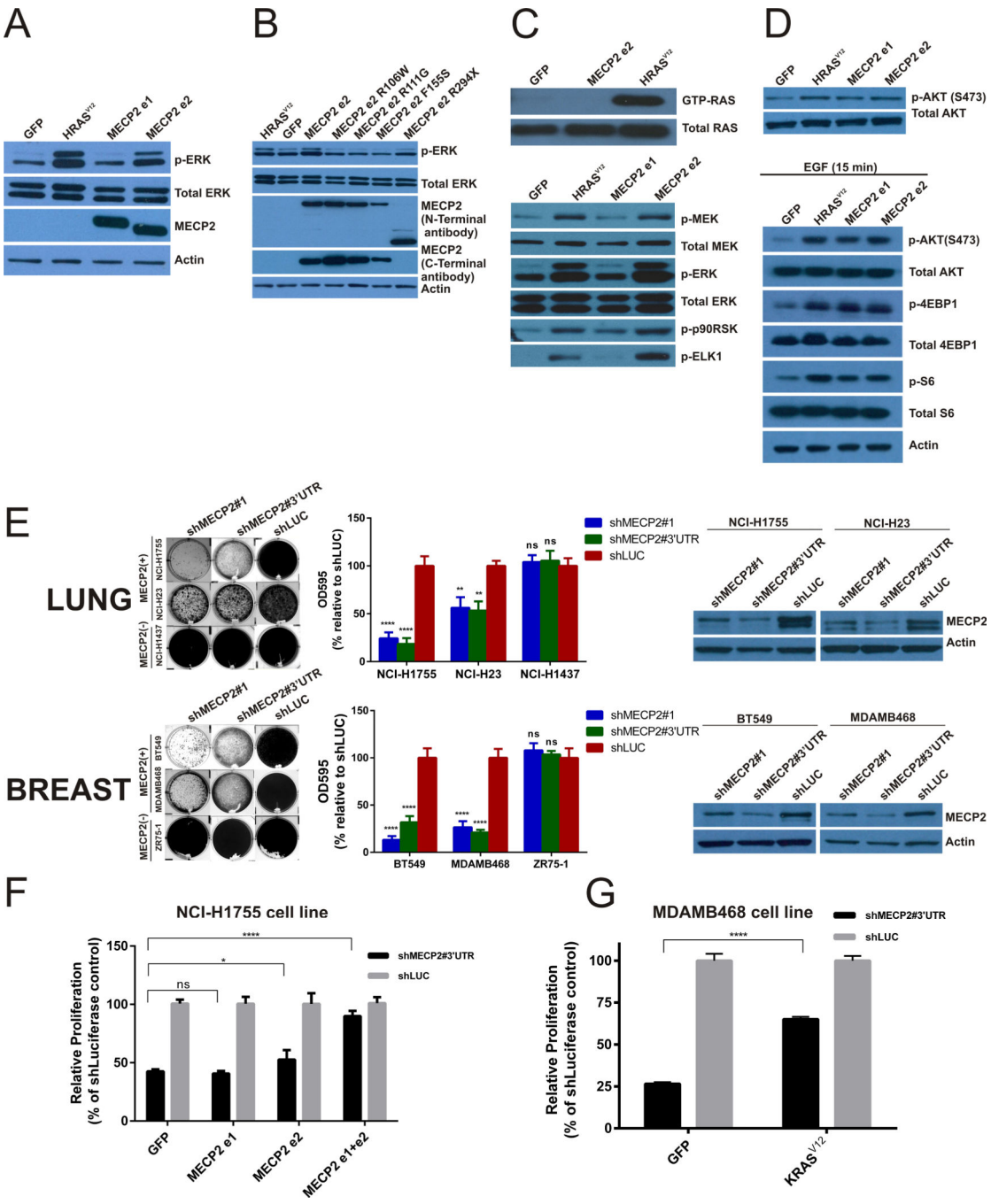


Figure 4. The *MECP2* Short Isoform e2 Activates The MAPK Pathway, Both *MECP2* Isoforms Activate the PI3K Pathway, and Human Cancer Cell Lines with Amplified *MECP2* are Addicted to *MECP2* Expression

(A, B) N minus RAS HMECs expressing the indicated genes were switched to minimal media without growth factors for five minutes (A), or twenty-four hours (B), and whole cell lysates were used for western blotting with the antibodies shown.

(C) Top Panel: Beads coupled to the ras binding domain (RBD) of CRAF were used to affinity precipitate GTP-bound RAS from cell lysates, and the precipitate was subject to immunoblotting with an antibody detecting all RAS proteins, labeled GTP-RAS; total cell

lysate was used with the same antibody and are labeled Total RAS. Bottom Panel: N minus RAS HMECs expressing the indicated genes were switched to minimal media without growth factors for five minutes, and whole cell lysates were used for western blotting with the antibodies shown.

(D) Top Panel: N minus RAS HMECs expressing the indicated genes were switched to minimal media without growth factors for five minutes, and then whole cell lysates were used for western blotting with the antibodies shown. Bottom Panel: N minus RAS HMECs expressing the indicated genes were switched to minimal media without growth factors for four hours, treated with 2.5 ng/ml EGF for 15 minutes, and then whole cell lysates were used for western blotting with the antibodies shown.

(E) Upper: Two lung adenocarcinoma cell lines with *MECP2* copy number gain and *MECP2*-overexpression, NCI-H1755 and NCI-H23, and the *MECP2* non-overexpressing lung adenocarcinoma cell line NCI-H1437 (for MECP2 protein expression in these lines, see Figure 1D) were infected with vectors encoding two different shRNAs targeting *MECP2* or shRNA targeting Luciferase as a control. Crystal violet staining was done 12 days post-infection, photographed and quantitated by OD595. Mean ±SD data from 3 independent viral infections (n =3) were analyzed using one-way ANOVA followed by post hoc Dunnett's multiple comparison test. Asterisks indicate statistically significant differences between shMECP2-infected cells and shLuciferase-infected cells from lung cancer lines: NCI-H1755 (****: $p<0.0001$ for shMECP2#1, and $p<0.0001$ for shMECP2#3'UTR), NCI-H23 (**: $p=0.0019$ for shMECP2#1, and $p=0.0013$ for shMECP2#3'UTR), NCI-H1437 (ns: non-significant, $p=0.7948$ for shMECP2#1 and $p=0.6709$ for shMECP2#3'UTR). Immunoblots for MECP2 show the effects of the indicated shRNAs.

Lower: Two triple negative breast cancer cell lines with *MECP2* copy number gain and *MECP2*-overexpression, BT549 and MDA-MB468, and *MECP2* non-overexpressing cell line ZR75-1 (for MECP2 protein expression in these lines, see Figure 1D) were infected with vectors encoding two different shRNAs targeting *MECP2* or shRNA targeting Luciferase as a control. Crystal violet staining was done 12 days post-infection, photographed (above) and quantitated (below) by OD595. Mean ±SD data from 3 independent viral infections (n =3) were analyzed using one-way ANOVA followed by post hoc Dunnett's multiple comparison test. Asterisks indicate statistically significant differences between shMECP2-infected cells and shLuciferase-infected cells from breast cancer lines: BT549 (****: $p<0.0001$ for shMECP2#1, and $p<0.0001$ for shMECP2#3'UTR), MDAMB468 (****: $p<0.0001$ for shMECP2#1, and $p<0.0001$ for shMECP2#3'UTR), ZR75-1 (ns: non-significant, $p=0.3928$ for shMECP2#1 and $p=0.7986$ for shMECP2#3'UTR). Immunoblots for MECP2 show the effects of the indicated shRNAs.

(F) NCI-H1755, an MECP2-overexpressing lung adenocarcinoma cell line used in 7A above, was infected with lentiviruses expressing the genes shown on the X axis, and then infected with the shMECP2#3'UTR (used in figure 4E above), or a control shRNA directed against luciferase. Cell growth was quantitated after 7 days by crystal violet (0.1%) staining of adherent cells followed by measurement of optical density of cell-associated dye extracted. Mean ±SD data from 3 independent viral infections (n =3) were analyzed using one-way ANOVA followed by post hoc Dunnett's multiple comparison test. Asterisks indicate statistically significant differences between GFP- versus *MECP2*-mediated rescue

of *MECP2* knockdown (****: p<0.0001 for MECP2 e1+e2; *: p=0.0340 for MECP2 e2; ns: non-significant, p=0.8970 for MECP2 e1).

(G) The *MECP2*-addicted triple negative breast cancer cell line MDAMB468 was infected with lentiviruses expressing either GFP or activated RAS ($KRAS^{G12V}$), and then infected with an shRNA targeting the MECP2 3'UTR (the same as used in figures 4E and 4F) or an shRNA targeting Luciferase as a control. Cell growth was quantitated after 7 days by crystal violet (0.1%) staining of adherent cells followed by measurement of optical density of cell-associated dye extracted. Mean ±SD data from 3 independent viral infections (n =3) were analyzed using unpaired two-tailed Student's t test. Asterisks indicate statistically significant difference between GFP- versus $KRAS^{V12}$-mediated rescue of MECP2 knockdown (****: p <0.0001).