Rett syndrome (RTT) is a prototype childhood neurological disease characterized by features that are observed in many other disorders ranging from autism to Parkinson’s disease and dystonia. The disorder affects ~1 in 10,000 females and is most often caused by mutations in the gene encoding methyl-CpG-binding protein 2 (MeCP2), a transcriptional regulatory protein. It has been shown that many of the features of RTT are reversible in mice (Gadalla et al., 2011; Guy et al., 2007), and that these features are probably due to dysfunction of neurons and supporting cells, rather than neural degeneration (Armstrong, 2002). These findings provide hope that some and perhaps most symptoms can be reversed in affected individuals if we discover effective therapies that can overcome the consequences of loss of function or dysfunction of MeCP2.

In September of 2011, the National Institute of Neurological Disorders and Stroke (NINDS), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the International Rett Syndrome Foundation (IRSF) and the Rett Syndrome Research Trust (RSRT) convened a workshop involving a broad cross-section of basic scientists, clinicians and representatives from the National Institutes of Health (NIH), the US Food and Drug Administration (FDA), the pharmaceutical industry and private foundations to assess the state of the art in animal studies of Rett syndrome (RTT). The aim of the workshop was to identify crucial knowledge gaps and to suggest scientific priorities and best practices for the use of animal models in preclinical evaluation of potential new RTT therapeutics. This review summarizes outcomes from the workshop and extensive follow-up discussions among participants, and includes: (1) a comprehensive summary of the physiological and behavioral phenotypes of RTT mouse models to date, and areas in which further phenotypic analyses are required to enhance the utility of these models for translational studies; (2) discussion of the impact of genetic differences among mouse models, and methodological differences among laboratories, on the expression and analysis, respectively, of phenotypic traits; and (3) definitions of the standards that the community of RTT researchers can implement for rigorous preclinical study design and transparent reporting to ensure that decisions to initiate costly clinical trials are grounded in reliable preclinical data.

Introduction

Rett syndrome (RTT) is a prototype childhood neurological disease characterized by features that are observed in many other disorders ranging from autism to Parkinson's disease and dystonia. The disorder affects ~1 in 10,000 females and is most often caused by mutations in the gene encoding methyl-CpG-binding protein 2 (MeCP2), a transcriptional regulatory protein. It has been shown that many of the features of RTT are reversible in mice (Gadalla et al., 2011; Guy et al., 2007), and that these features are probably due to dysfunction of neurons and supporting cells, rather than neural degeneration (Armstrong, 2002). These findings provide hope that some and perhaps most symptoms can be reversed in affected individuals if we discover effective therapies that can overcome the consequences of loss of function or dysfunction of MeCP2.
There is a crucial need to develop effective treatments for RTT. Encouragingly, several key findings suggest that RTT could be treatable in humans, and might be a promising model for developing rigorous paradigms for evaluating therapeutic interventions not only in RTT but in other postnatal childhood disorders as well. First, RTT typically manifests months after birth, arguing that key embryonic and perinatal developmental steps take place normally in affected individuals. Second, there are several excellent mouse models in which many of the somatic, behavioral and physiological changes observed in individuals with RTT are reproduced (discussed in detail below). Third, the findings that some RTT-like symptoms are reversible in mouse models following reactivation of silent MeCP2 alleles (Guy et al., 2007; Lioy et al., 2011; Robinson et al., 2012) or transgene-mediated MeCP2 replacement (Alvarez-Saavedra et al., 2007; Collins et al., 2004; Giacometti et al., 2007; Jugloff et al., 2008; Luikenhuis et al., 2004), and that most RTT-like symptoms are reproduced following loss of MeCP2 in adult animals (Cheval et al., 2012; McGraw et al., 2011; Nguyen et al., 2012), argue that most of the disease phenotypes are caused by functional disturbances of neural circuits, rather than irreversible developmental brain abnormalities. Thus, the challenge for the field now is to build on these exciting findings by identifying therapeutic opportunities, testing them rigorously in RTT models to determine their effectiveness in improving clinically relevant outcome measures, and establishing their safety with chronic use.

Given the emergence of new therapeutic leads in the field (cf. Abdala et al., 2010; De Filippis et al., 2012; Deogracias et al., 2012; Kron et al., 2012; McCauley et al., 2011; Nag and Berger-Sweeney, 2007; Ogier et al., 2007; Roux et al., 2007; Schmid et al., 2012; Tropea et al., 2009; Zanella et al., 2008), RTT models hold great promise for translational research, particularly for prioritizing and validating potential treatment strategies prior to launching costly clinical trials. Unfortunately, as discussed later in this review, there is a long history of failure in translating promising findings from preclinical animal models to clinical success, especially for neurological disorders. Therefore, it is crucial that the RTT research community develops best practices and standards for performing preclinical trials, and identifies the best path forward for justifying the advancement of preclinical discoveries to clinical trials. These steps are essential to avoid involving individuals with RTT in unnecessary clinical trials, wasting precious research money, and raising unwarranted expectations for the patients and their families.

With these goals in mind, the National Institute of Neurological Disorders and Stroke (NINDS), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), and private United States organizations supporting RTT research [the International Rett Syndrome Foundation (IRSF) and the Rett Syndrome Research Trust (RSRT)] held a workshop entitled ‘Setting Priorities for Therapy Development in Rett Syndrome’ at the Hyatt Regency Bethesda (Bethesda, MD) on 25-27 September 2011 to discuss how to optimize the predictive value of animal models in RTT preclinical research and avoid the pitfalls that often lead to failure on clinical translation. In addition to these funding agencies and foundations, workshop participants included members of the RTT research community and the pharmaceutical industry, clinicians, and representatives from the US Food and Drug Administration (FDA). The workshop agenda and participant list are available in supplementary material Tables S1 and S2, respectively.

A key component of the workshop was the creation of working groups that were responsible for reviewing the state of the art in phenotypic analyses of mouse models of RTT. Four groups were created, each focusing on different phenotypic domains: (1) general health, locomotion and lifespan (led by Dr James Eubanks, University of Toronto); (2) respiratory and autonomic control (led by Drs David Katz, Case Western Reserve University School of Medicine, and Jeffrey Neul, Baylor College of Medicine); (3) cognitive, social and anxiety phenotypes (led by Dr Joanne Berger-Sweeney, Tufts University); and (4) cellular and synaptic phenotypes (led by Dr Lucas Pozzo-Miller, The University of Alabama at Birmingham). Reporting on key outcomes from the workshop, as well as subsequent discussions among the working groups, this paper reviews the phenotypic characteristics of currently available mouse models of RTT, highlights current knowledge gaps, and identifies criteria and best practices for preclinical study design that we hope will optimize the ability of the RTT research community to translate basic findings into new therapeutic approaches. We begin with a brief overview of the principal clinical features of RTT, which is a necessary foundation for evaluating the degree to which existing mouse models reproduce phenotypes of the human disease, and for identifying sensitive and relevant outcome measures for preclinical and clinical trials.

**Genetic and clinical-pathological features of RTT**

Loss-of-function mutations in MECP2, an X-linked gene, account for the vast majority (~95%) of typical RTT cases (Amir et al., 1999). Most mutations arise spontaneously (de novo) in the paternal germ line; thus, individuals with RTT are typically females who, owing to X-chromosome inactivation, are somatic mosaics for normal and mutant MECP2. Boys with mutations that cause RTT in females typically die before or soon after birth with a severe encephalopathy (discussed in more detail below).

RTT is distinguished by its unique time course and phenotypic complexity. Affected individuals present with postnatal neurological regression, usually starting between 1.5 and 3 years of age (but sometimes as early as 6 months of age), with loss of acquired hand skills and spoken language and, in some cases, social withdrawal or extreme irritability that can resemble autism (Hagberg, 2002; Neul et al., 2010). After regression, there is a stabilization of skills, rather than a relentless progression, a feature that differentiates RTT from neurodegenerative conditions such as Batten disease or Huntington’s disease. During this pseudo-stationary or plateau stage, characteristic features of RTT such as repetitive hand movements (stereotypes), which can be present before or during regression, become more prominent. Later in life, many affected individuals enter a stage of motor decline in which ambulation can be lost, and Parkinsonian features such as rigidity and hypomimia become prominent (FitzGerald et al., 1990a; FitzGerald et al., 1990b). Mutations in other genes such as cyclin-dependent kinase like 5 (CDKL5) and forkhead box G1 (FOXG1) can cause phenotypes overlapping with those seen in RTT (Archer et al., 2006; Ariani et al., 2008); however, several features, such as congenital onset and infantile spasms in CDKL5-mutant patients, and congenital onset and hypoplasia of the corpus callosum in FOXG1-mutant patients, distinguish these disorders from typical RTT (Kortüm et al., 2011).

During the regression stage, some individuals with RTT develop autistic features that include social withdrawal, avoidance of eye contact and indifference to visual or auditory stimuli (Mount et al., 2002a; Mount et al., 2003). After regression, some of these autistic features decrease, and most affected individuals develop intense eye gaze that they use for communication (Coenraads, 2007; Kaufmann et al., 2012). Recent work has shown that features such as stereotypes and lack of...
language skills persist throughout the life of affected individuals, although hand stereotypes can change from rapid movements to midline hand clapping with age. Additional behavioral problems include anxiety in response to novel situations (Mount et al., 2002b), increased behavioral rigidity and increased pain tolerance (Downs et al., 2010). Individuals with RTT are considered to have severe intellectual disability; however, because affected individuals have severe impairments in their ability to communicate, it is difficult to make accurate assessments of their intellectual ability (Baptista et al., 2006; Neul et al., 2010).

Movement abnormalities are a major issue in RTT (FitzGerald et al., 1990a; FitzGerald et al., 1990b), with the most obvious being the repetitive hand stereotypes, which seem to interfere with volitional hand use. Gait is almost always disrupted, with evidence of ataxia and apraxia. Dystonia is common, seen first in the ankles and eventually progressing to many joints. Axial hypotonia is present early in the disease course but, as children become young adults, increased tone with features of rigidity becomes more prominent. Additional movement abnormalities include tremor, myoclonus, chorea, facial grimacing and severe teeth grinding. Most individuals with RTT have scoliosis, and some require surgical intervention (Percy et al., 2010).

Nutrition and gastrointestinal function are also major clinical issues in RTT, and there is marked growth failure in most affected individuals (Tarquinio et al., 2012). It has long been recognized that head growth is impaired, resulting in acquired microcephaly (Hagberg et al., 1983), and height and weight are usually markedly diminished (Schultz et al., 1993). However, a subset of individuals with RTT are overweight or obese (Renieri et al., 2009), a feature that is often associated with higher functioning and possibly improved oromotor skills (Motil et al., 1999). Many individuals with RTT have various gastrointestinal problems, including significant chewing and swallowing difficulties, gastroesophageal reflux, gastrointestinal dysmotility and severe constipation, which severely decrease the quality of life for patients and their families (Motil et al., 2012).

Dysregulation of breathing and autonomic homeostasis are very common in RTT. Respiratory abnormalities, which include periods of forceful breathing (hyperventilation), severe pauses in breathing (including breath holds) that can cause cyanosis and even loss of consciousness, and abnormal cardiorespiratory coupling, are more severe during wakefulness than during sleep (Elian and Rudolf, 1991; Julu et al., 2001; Julu and Witt Engerström, 2005; Marcus et al., 1994; Weese-Mayer et al., 2008; Weese-Mayer et al., 2006) and can be exaggerated during periods of excitement or stress. Autonomic abnormalities include periods of vasomotor disturbance (usually associated with cold hands and feet), abnormal sweating, decreased heart rate variability, evidence of sympathetic-parasympathetic imbalance and prolongation of corrected QT interval (an indication of abnormal cardiac electrical activity) in a subset of individuals (Guidieri et al., 2004; McCauley et al., 2011; Sekul et al., 1994). One quarter of deaths in RTT are sudden and unexpected (Kerr et al., 1997), and might result from complications of cardiorespiratory dysfunction.

Brain electrical activity is not typical in individuals with RTT, as shown by the markedly disrupted pattern observed on electroencephalograms (EEGs) (Glaze et al., 1998) and the high probability of seizures (Glaze et al., 2010). Seizures, ranging from complex partial to generalized tonic-clonic, are most commonly seen after other symptoms appear (usually after age two) and correlate with the severity of the phenotype. In addition to true epileptic events, individuals with RTT also have non-epileptic paroxysmal events, and video EEG is needed to differentiate between them (Glaze et al., 1998).

Despite the severity and phenotypic complexity of RTT, the brains of individuals with RTT do not show gross neuropathological changes, nor evidence of neuronal or glial atrophy, degeneration, gliosis, or demyelination, indicating that RTT is not a neurodegenerative disorder (Jellinger et al., 1988; Reiss et al., 1993). Smaller total brain volume and smaller neurons (but with a higher cell density) have been observed in several brain regions, including the cerebral cortex, hypothalamus and the hippocampal formation (Bauman et al., 1995a; Bauman et al., 1995b). The size and complexity of dendritic trees are reduced in cortical pyramidal cells (Armstrong et al., 1995; Armstrong et al., 1998), and levels of microtubule-associated protein-2 (MAP-2), a protein involved in microtubule stabilization, are lower throughout the neocortex of RTT autopsy material (Kaufmann et al., 2000; Kaufmann et al., 1995). In addition, the density of dendritic spines is lower in pyramidal neurons of the frontal cortex (Belichenko et al., 1994; Jellinger et al., 1988) and in the CA1 region of the hippocampus (Chapleau et al., 2009).

### Mouse models of RTT

Identification of MECP2 as the disease-causing gene led rapidly to the development of mouse models of RTT (Table 1) that recapitulate, to varying degrees, the underlying molecular and genetic defects and symptoms of the human disease (Table 2). Established models include mice carrying either global alleles (null, hypomorphic, large deletions and point mutations) or conditional null alleles (reviewed in Calfa et al., 2011b). As discussed below, a major focus at the workshop concerned the utility of the various models for translational studies. One point of consensus was that conditional alleles (i.e. cell-specific deletions created by cross-breeding of floxed Mecp2 mice with Cre-deleter lines), although useful for elucidating mechanisms of neurological dysfunction, are unlikely to be of value in translational...
studies because, by design, they do not recapitulate the global loss of 
MeCP2 that is characteristic of RTT.

Global Mecp2 alleles generally fall into three categories: (1) null 
mutations (MeCP2tm1.1Bird, MeCP2tm1.1Irsf or MeCP2tm1.1Meg) (Chen et al., 
2001; Guy et al., 2001; Pelka et al., 2006), which are similar to the large 
deletions found in ~10% of affected people; (2) truncations and single-
nucleotide mutations that approximate or reproduce mutations 
found in individuals with RTT (codon 308-C-terminal truncation: 
MeCP2tm1.1Hao; R168X: MeCP2tm1.1Coyle and MeCP2tm1.1Hup; T158A: 
MeCP2tm1.1Joez; and R255X: MeCP2tm1.1Irsf) (Brendel et al., 2011; Goffin 
et al., 2012; Jentarra et al., 2010; Lawson-Yuen et al., 2007; Shahbazian 
et al., 2002); and (3) point mutations that either model related 
neurodevelopmental disorders [A140V: MeCP2tm1.1Vnar (Jentarra et al., 
2010)] or provide mechanistic insight into MeCP2 function [S421A: 
MeCP2tm1.1Meg (Cohen et al., 2011)]. Mutant mice that are hypomorphic 
for MeCP2 (MeCP2tm1.1Bird) (Kerr et al., 2008; Samaco et al., 2008) have 
also been created; these animals express MeCP2 at ~50% of wild-type 
levels due to retention of a selectable marker cassette in the floxed MeCP2 
allele and, as expected, exhibit more modest behavioral impairments 
compared with the complete nulls. Phenotypic details of the various 
RTT mouse models are summarized in supplementary material 
Tables S3-S7.

During the workshop, discussion of mouse models focused on three 
key concepts used to evaluate the utility of a model for translational 
studies: (1) construct validity (similarities in underlying molecular 
mechanisms in humans and mice); (2) reproducibility (findings 
replicated in more than one laboratory) and robustness (effect size, and 
whether or not findings are generalizable to more than one model 
and/or experimental condition); and (3) face validity (similarities in 
anatomical, physiological and/or behavioral phenotypes in humans and 
mice).

**Construct validity**

To some degree, all global null or hypomorphic alleles have construct 
validity, given that most disease-causing mutations in human RTT cause 
loss of MeCP2 function (e.g. Kudo et al., 2001). However, as noted above, 
only 10% of mutations in individuals with RTT are deletions that are 
large enough to approximate a true null condition. Moreover, although 
male mice have, until recently, been used most frequently in preclinical 
studies of RTT, male nulls and hypomorphs do not exhibit the 
genetic mosaicism that is characteristic of the vast majority of 
individuals with RTT, who are female. Furthermore, it is not known 
whether complete loss of MeCP2 in cells (either in hemizygous males 
or heterozygous females) recapitulates the effects of a point mutation 
or truncated protein, including the possibility of toxic gain of function 
with some of the human mutations. Thus, translational studies in mice 
carrying null alleles might only be relevant to a relatively small subset 
of individuals with RTT. However, many MECP2 point mutations, in 
addition to large deletions, are believed to eliminate gene function; in 
particular, some mutations in the methyl-binding and transcriptional 
repression domains cause the most severe RTT phenotypes and might 
be relevant to the null allele. Therefore, current attempts to establish 
the construct validity of various RTT mouse models (e.g. null versus 
point mutations) would benefit from detailed comparisons of how the 
corresponding alleles, when studied in the same genetic background, 
alter target gene expression, cell physiology and neurological function. 
It would also be helpful to know whether the two most commonly used 
null alleles, MeCP2tm1.1Bird and MeCP2tm1.1Irsf, are functionally equivalent, 
given that they were generated using targeting strategies directed against 
slightly different sequences in the coding region of MeCP2. Ultimately, 
choosing mouse models for translational studies might require 
compromises between construct validity and practical concerns – i.e. 
to optimize phenotypes that are robust and that develop within a 
timeframe that is cost effective for large-scale preclinical studies.

### Table 2. Comparison of phenotypes in RTT individuals and MeCP2-mutant mice

<table>
<thead>
<tr>
<th>Core phenotypes in RTT individuals</th>
<th>Similar phenotypes in MeCP2-mutant mice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphological</td>
<td>++</td>
</tr>
<tr>
<td>Neuronal hypotrophy (reduced</td>
<td>++</td>
</tr>
<tr>
<td>neuronal soma size, smaller</td>
<td></td>
</tr>
<tr>
<td>dendritic arbors and lower spine</td>
<td></td>
</tr>
<tr>
<td>density)</td>
<td></td>
</tr>
<tr>
<td>Respiratory and autonomic control</td>
<td>++</td>
</tr>
<tr>
<td>Abnormal breathing (irregular</td>
<td></td>
</tr>
<tr>
<td>pattern, respiratory pauses,</td>
<td></td>
</tr>
<tr>
<td>increased mean frequency)</td>
<td></td>
</tr>
<tr>
<td>Prolonged QTCc</td>
<td>+</td>
</tr>
<tr>
<td>Vasomotor disturbances</td>
<td>nd</td>
</tr>
<tr>
<td>Gastroduodenal dysmotility</td>
<td>nd</td>
</tr>
<tr>
<td>Motor function</td>
<td>++</td>
</tr>
<tr>
<td>Gait abnormalities</td>
<td></td>
</tr>
<tr>
<td>Stereotypies</td>
<td>+</td>
</tr>
<tr>
<td>Tremors</td>
<td>++</td>
</tr>
<tr>
<td>Early hypotonia</td>
<td>nd</td>
</tr>
<tr>
<td>Lack of purposeful hand movements</td>
<td>? (poor nest building might reflect</td>
</tr>
<tr>
<td></td>
<td>decreased forepaw function)</td>
</tr>
<tr>
<td>Parkinsonian features</td>
<td>? (mice exhibit hypokinesis; mechanism</td>
</tr>
<tr>
<td></td>
<td>unknown)</td>
</tr>
<tr>
<td>Dystonia</td>
<td>? (hindlimb clasp might be a sign of</td>
</tr>
<tr>
<td></td>
<td>dystonia)</td>
</tr>
<tr>
<td>Cognition and behavior</td>
<td>++</td>
</tr>
<tr>
<td>Cognitive deficits</td>
<td></td>
</tr>
<tr>
<td>Social withdrawal</td>
<td>+/-</td>
</tr>
<tr>
<td>Increased anxiety</td>
<td>+/-</td>
</tr>
<tr>
<td>Loss of speech</td>
<td>?</td>
</tr>
<tr>
<td>Repetitive behaviors</td>
<td>? (some indication of repetitive and/or</td>
</tr>
<tr>
<td></td>
<td>excessive grooming)</td>
</tr>
<tr>
<td>Other</td>
<td>++</td>
</tr>
<tr>
<td>Reduced lifespan</td>
<td></td>
</tr>
<tr>
<td>EEG abnormalities and seizures</td>
<td>++</td>
</tr>
<tr>
<td>Neurological regression</td>
<td>++</td>
</tr>
<tr>
<td>Sleep cycle disturbances</td>
<td>nd</td>
</tr>
<tr>
<td>Scoliosis</td>
<td></td>
</tr>
</tbody>
</table>

++, phenotypes similar to one mouse model; +, phenotypes similar to human and reported in only one mouse model to date; +/-, conflicting data in different mouse models; ?, suggestive or inconclusive data; nd, not determined. Detailed mouse phenotyping data are available in online supplementary material Tables S3-S6.
What is notable in RTT mouse models is evidence of construct validity beyond the genetic mutations themselves, including neurochemical abnormalities found in individuals with RTT. For example, diverse mouse models reproduce the reduced levels of biogenic amines and brain-derived neurotrophic factor (BDNF) that are found in the brains of RTT patients (Wang et al., 2006; Deng et al., 2007; Li et al., 2012; Chang et al., 2006). Thus, removing MeCP2 function from mice reproduces at least some of the same cell signaling deficits as MECP2 mutations in humans.

Reproducibility and robustness

In addition to the allelic diversity of current RTT models, RTT research is further confounded by the fact that different laboratories study different alleles on different genetic backgrounds and at different developmental stages, making it difficult to discern the effects of a specific allele from genetic, environmental and maturational influences. Expression of the same Mecp2 allele on different genetic backgrounds can confer significant differences in phenotypic effects, including but not limited to time of onset and symptom severity. Similarly, the degree to which animal husbandry practices (including food, water, lighting, noise, handling and bedding), maternal age and quality of maternal care influence the phenotypic effects of Mecp2 alleles has not been studied. Furthermore, disease progression across the lifespan has not been carefully evaluated in most models, especially in heterozygous females, as discussed below. For translational studies, therefore, the variations in phenotype that are influenced by genetic background and animal husbandry practices highlight the importance of validating phenotypes and therapeutic efficacy in multiple models and in multiple laboratories. Such validation will hopefully mitigate the possibility that a given study is treating idiosyncratic gene-gene or gene-environment interactions that might not generalize or be relevant to the human disease.

Fortunately, and as would be expected based on the high penetrance of MECP2 mutations in humans, at least some phenotypes that mimic symptoms of human RTT are common to multiple Mecp2-mutant mouse models (see below). However, relatively little is known about whether RTT mouse models recapitulate the unique features of progression, regression and stabilization that define the human disease. This might reflect the overall bias of the scientific community, which has performed most phenotypic analyses in male hemizygous mice. These animals might best model the small number of human male patients who are hemizygous for mutant MECP2 and present, not with clinically defined RTT, but with a severe encephalopathy accompanied by a failure to acquire skills, rather than regression. An increasing number of laboratories are shifting to phenotypic analysis of heterozygous female mice, which might yield models that more closely phenocopy human RTT, including temporal features such as disease regression.

Face validity: RTT phenotypes recapitulated in mouse models

The following sections provide an overview of the key pathophysiological findings that have been reported in RTT mouse models carrying global (rather than cell-specific) Mecp2 mutant alleles. Most models exhibit a broad spectrum of phenotypes that are similar to those seen in RTT individuals, including shortened lifespan, motor and sensory impairments, breathing abnormalities, cognitive and behavioral dysfunction, and cellular and synaptic defects (reviewed in Calfa et al., 2011b) (summarized in Table 2). However, care must be taken not to assume, on the basis of face validity alone, that similar behavioral phenotypes in mice and humans necessarily arise from the same underlying pathophysiological mechanisms. Emphasis here is given to those findings that are most robust and reproducible across different models and/or genetic backgrounds and in different laboratories, with the goal of identifying the models and phenotypes that will be most useful for translational studies. This overview is by no means exhaustive; further details and additional references for each model are provided in the accompanying online supplemental material (supplementary material Tables S3-S7). Male mice that are hemizygous for Mecp2 null alleles are referred to as ‘Nulls’ and females that are heterozygous for Mecp2 null alleles are referred to as ‘Hets’.

General health, locomotion and lifespan

Gross motor dysfunction and shortened lifespan are common features of diverse RTT mouse models on different genetic backgrounds and broadly recapitulate the clinical presentation in individuals with RTT (Table 2; supplementary material Table S3). Mecp2-mutant mice exhibit robust and reproducible abnormalities in motor behaviors, including hypoactivity (Chen et al., 2001; Guy et al., 2001; Jugloff et al., 2008; Stearns et al., 2007), impaired balance and coordination (Alvarez-Saavedra et al., 2007; Jugloff et al., 2008; Kondo et al., 2008; Shahbazian et al., 2002; Stearns et al., 2007), spontaneous tremors (Alvarez-Saavedra et al., 2007; Brendel et al., 2011; Chen et al., 2001; Goffin et al., 2012; Guy et al., 2001; Lawson-Yuen et al., 2007; Pelka et al., 2006; Shahbazian et al., 2002; Stearns et al., 2007), hindlimb clasping (Brendel et al., 2011; Chen et al., 2001; Goffin et al., 2012; Guy et al., 2007; Guy et al., 2001; Kondo et al., 2008; Lawson-Yuen et al., 2007; Pelka et al., 2006; Stearns et al., 2007), and impaired limb and postural reflexes (Picker et al., 2006; Santos et al., 2007). In male Nulls, these deficits occur early (within 6 weeks of birth) and are generally more pronounced than in female Hets; female Hets typically develop milder motor impairments (usually not seen before 10 weeks of age) and exhibit greater phenotypic variability (Stearns et al., 2007). In addition, excessive and repetitive grooming has been observed in male Mecp2<sup>tm1.1Jae</sup> Nulls and might represent a form of motor stereotypy (Stearns et al., 2007). To a first approximation, many of these phenotypes seem to mimic motor deficits exhibited by individuals with RTT. For example, hindlimb clasping, although not specific to RTT mice, might represent a form of motor dysfunction. However, further work is required to determine the degree to which these human and mouse phenotypes truly share common underlying neurological mechanisms.

Lifespan is severely shortened in male Nulls (including both the commonly employed Mecp2<sup>tm1.1Bir</sup> and Mecp2<sup>tm1.1Bae</sup> models), which rarely live longer than 3 months. Prior to death, male Nulls often exhibit severe hypoactivity, kyphosis and disheveled fur, and typically undergo severe weight loss. In female Hets carrying null alleles, there is a higher than normal rate of sudden and unexpected death (unpublished observations from multiple groups), although kyphosis and disheveled fur are less reproducibly observed. Lifespan is also reduced in most of the male hemizygous human (targeted) mutation models (Goffin et al., 2012; Lawson-Yuen et al., 2007; Shahbazian et al., 2002), but the degree varies with the specific Mecp2 allele and/or genetic background. An exception is the Mecp2<sup>tm1.1Vnar</sup> (A140V) mutant line, which has an apparently normal lifespan and, other than cellular abnormalities, does not exhibit many obvious phenotypic defects (Jentarra et al., 2010); these findings are consistent with clinical data showing that missense A140V
mutations are usually associated with milder phenotypes in males and have not been reported to cause classical RTT in females.

In contrast to the relatively robust phenotypes described above, genetic background seems to have a significant impact on body weight. Mecp2 null alleles on a C57BL/6 background tend to have reduced body weight (Guy et al., 2001; Pelka et al., 2006; Stearns et al., 2007; Ward et al., 2011), whereas alleles on the 129 background often exhibit increased body weight (Chen et al., 2001; Shahbazian et al., 2002). Lonetti and colleagues used a mixed genetic background (B6.129SF1) and found no evidence of increased body weight (Lonetti et al., 2010).

Respiratory and autonomic phenotypes

The most robust breathing phenotype in RTT mouse models that has been reproducibly observed by different laboratories and in different mouse strains is abnormal variation in respiratory cycle length in room air, including respiratory pauses and periods of tachypnea associated with decreased expiratory time and increased mean breathing frequency (Table 2; supplementary material Table S4). These breathing abnormalities closely phenocopy respiratory dysfunction observed in individuals with RTT. Respiratory pauses occur in male Nulls, and female Hets and hypomorphs, with an earlier onset in Nulls (by ~5 weeks) compared with Hets (~10 weeks). Null males and Het females also exhibit exaggerated respiratory reflexes, including enhanced vagally induced respiratory pauses and increased hypoxic ventilatory responses (Bissonnette and Knopp, 2006; Johnson et al., 2012; Roux et al., 2008; Stettner et al., 2007; Voituron et al., 2009).

Although cardiac phenotypes have been less well characterized, recent work indicates that male Nulls and female Hets exhibit prolongation of the corrected QT interval and an increased susceptibility to induced cardiac arrhythmia and cardiac death. This seems to be a progressive phenotype as female Hets exhibit these phenotypes at 11 but not 4 months of age (Mccauley et al., 2011).

Cognitive, social and anxiety phenotypes

As noted above, the severity of communication deficits in individuals with RTT has hampered systematic evaluation of their cognitive abilities. Thus, although cognitive deficits have been identified in RTT mouse models (Table 2; supplementary material Table S5), the face validity of these findings has been difficult to establish. One of the most reliable and robust cognitive phenotypes in RTT mice – i.e. replicated by different laboratories, in different RTT mouse models with diverse Mecp2 alleles and/or genetic backgrounds – is impairment in contextual fear conditioning, which is a test of associative learning and memory. This phenotype is apparent in male Nulls by about 6 weeks of age (Stearns et al., 2007); similarly, the human mutation models tested to date – male Mecp2tm1 Hzoe and Mecp2tm1Hzo mice – show impairments in contextual fear conditioning by 10 and 20 weeks of age, respectively. In contrast and similar to the severe motor phenotype, male Mecp2tm1.1Joezand mice showed a clear deficit in contextual fear conditioning (Lonetti et al., 2010; Kerr et al., 2012; Pelka et al., 2006; Stearns et al., 2007). The findings are the opposite in male Mecp2tm1Hzo mutants, which are more anxious in the same tasks (Moretti et al., 2005; Moretti et al., 2006; Shahbazian et al., 2002). Thus, the impact of reduced MeCP2 function per se on social behavior remains very unclear, as is the relevance of the mouse phenotypes identified thus far to the social withdrawal and autistic features seen in some individuals with RTT.

Likewise, measures of anxiety vary widely depending on allele, genetic background and testing paradigm. Using the elevated plus and zero mazes, male Nulls and Mecp2tm1.1Joez mice seem less anxious than wild types (i.e. they spend more time in light spaces of the maze) (Goffin et al., 2012; Kerr et al., 2012; Pelka et al., 2006; Stearns et al., 2007). The findings are the opposite in male Mecp2tm1Hzo mutants, which are more anxious in the same tasks (De Filippis et al., 2010; McGill et al., 2006). Using the open field, where mice ambulate between the center of the circular field (less anxious behavior) and the periphery of the field (more anxious behavior), male Mecp2tm1Hzo mutants and female Mecp2tm1.1Joez Hets seem more anxious than wild types (De Filippis et al., 2010; Lonetti et al., 2010).

A few laboratories have demonstrated that RTT mice exhibit abnormal ultrasonic vocalizations when pups are separated from their mothers during early postnatal development. Specifically, beginning at postnatal days 4–6, neonatal ultrasonic vocalizations are altered, although in opposite directions in the two Mecp2 models tested to date: vocalizations are decreased in male Mecp2tm1Hzoe mice and increased in both male Null and female Het Mecp2tm1.1Joez mice compared with wild-type littermates (De Filippis et al., 2010; Picker et al., 2006). Whether these apparently conflicting results reflect true strain differences remains to be determined. Furthermore, although many individuals with RTT lack speech, most are very vocal. A better understanding of the underlying neural circuit dysfunction might help to differentiate between the preservation of vocalization and the loss of speech in RTT. Clearly, this is an important area for further investigation, because loss of acquired speech is a hallmark of classical RTT, and identifying a robust mouse model for this feature would be a valuable tool for understanding the underlying pathophysiology and for future translational studies.

Importantly, many behavioral tasks, including all of the anxiety measures, require an animal to ambulate. Therefore, testing of behavioral outcomes is often confounded by the severe motor impairments that are characteristic of most RTT mouse models. Furthermore, anxiety measures are highly sensitive (more so than for...
most behavioral tasks) to prior handling, the order of testing and the laboratory environment. Thus, for behavioral tasks in general, and anxiety measures in particular, it is important to control carefully for these variables.

Cellular, synaptic and circuit phenotypes
To some degree, the reproducibility and robustness of cellular and synaptic phenotypes in different Mecp2-mutant mice is difficult to evaluate owing to the methodological differences between studies published from different laboratories, and because not all available mouse models have been systematically analyzed. However, the prototypical features of human RTT neuropathology – such as smaller neurons, higher neuronal packing density, reduced dendritic arbors and abnormal dendritic spines – have all been found in models analyzed thus far (Table 2; supplementary material Table S6). In addition, studies of different strains of MeCP2-deficient mice have revealed consistent impairments in one or more physiological properties at the cellular and synaptic level in all brain regions studied so far (supplementary material Table S6). These impairments include network hyperexcitability in the brainstem and hippocampus (Calza et al., 2011a; Kline et al., 2010; Kron et al., 2012; Medrihan et al., 2008; Zhang et al., 2008), network hypoexcitability and synaptic hypoconnectivity in the cerebral cortex (Dani et al., 2005; Dani and Nelson, 2009), altered intrinsic neuronal electrical properties in the locus ceruleus and substantia nigra (Gantz et al., 2011; Taneja et al., 2009), and dysregulation of transmitter release in cultured hippocampal neurons and chromaffin cells (Nelson et al., 2006; Wang et al., 2006). However, these physiological impairments exhibit significant regional specificity: decreased excitatory synaptic drive is found in cortical circuits, whereas increased neuronal or synaptic excitability is found in the hippocampus and multiple brainstem regions involved in autonomic control (reviewed in Shepherd and Katz, 2011) (see also Kron et al., 2012). Modest impairments in individual elements (e.g. GABAergic synapses) might lead to different outcomes depending on the spatiotemporal construction and connectivity of the specific neuronal network. In addition, because most physiological studies at the cellular and synaptic level have been carried out in male Nulls, it has rarely been feasible to distinguish primary cell-autonomous deficits from the secondary or compensatory changes that occur in females with a mosaic pattern of neuronal and glial MeCP2 deficiency [see, however, Taneja et al. (Taneja et al., 2009)].

Several laboratories have documented impairments in long-term potentiation (LTP) and depression (LTD), which are two forms of synaptic plasticity associated with learning and memory, in brain slice preparations from Mecp2 mutants, including Mecp2<sup>tm1.1ls</sup>, Mecp2<sup>tm1.1ls</sup> and Mecp2<sup>tm1.1ho</sup> Null mice (reviewed in Boggio et al., 2010). Reduced LTP has been found at excitatory CA3-to-CA1 synapses in the hippocampus (Asaka et al., 2006) and in layer II/III of the primary somatosensory cortex (Lonetti et al., 2010). Reduced LTD has been demonstrated in area CA1 (Asaka et al., 2006; Moretti et al., 2006). So far, most studies of synaptic plasticity have analyzed populations of neurons and synapses using extracellular electrodes. Importantly, analysis of monosynaptic connections between pyramidal neurons in layer V of the primary somatosensory cortex of Mecp2<sup>tm1.1ho</sup> Null mice using intracellular electrodes demonstrated that LTD was intact at these synapses, provided that sufficient postsynaptic depolarization was achieved by step depolarization or by evoked action potentials during the induction of spike-timing-dependent plasticity (Dani and Nelson, 2009). Thus, it is not yet clear whether LTP phenotypes that are identified using extracellular electrodes actually reflect weak and sparse connections rather than specific deficits in classical mechanisms of LTP generation.

The presentation of seizure-like behaviors and atypical EEGs in RTT individuals [i.e. focal, multifocal and generalized epileptiform abnormalities, with rhythmic slow theta activity in the frontal-central regions (Glaze, 2005)] is recapitulated in Mecp2-mutant mice as a constellation of EEG abnormalities, with or without associated seizure-like behaviors. Indeed, abnormal EEGs have been described in male Null and female Het Mecp2<sup>tm1.1ls</sup> mice (D’Cruz et al., 2010; Goffin et al., 2012; Liao et al., 2012; Wither et al., 2012), as well as in Mecp2<sup>tm1.1ho</sup> Null mice (Shahbazian et al., 2002) and Mecp2<sup>tm1.1ho</sup> mutants (Goffin et al., 2012).

Improving the quality and rigor of preclinical research in RTT: a necessary step to achieving translational success
In addition to the importance of establishing well-characterized and validated animal models, the workshop participants discussed the need to improve the quality and rigor of preclinical studies as an important step in enhancing the predictive value of RTT translational research. Barriers to translational success, regardless of the disease in question, often include a lack of rigorous standards and transparency in reporting preclinical studies, as well as publication bias caused by the under-reporting of negative results in the scientific literature. At the workshop, Dr Shai Silberberg (NINDS) reviewed recently published meta-analyses showing that unintended biases in animal studies tend to substantially inflate effect sizes, or lead to Type I errors (false positives) and overestimations of potential therapeutic efficacy. These unintended biases include the lack of allocation concealment [i.e. the investigators have no knowledge of the experimental group to which an animal belongs (Fisher et al., 2009)], blinded assessment of outcome and random allocation of subjects to experimental groups (see Bebarta et al., 2003; Macleod et al., 2005; Ranshooff and Gourlay, 2010; Sene et al., 2007; Landis et al., 2012). Furthermore, in recent efforts to reproduce original findings of potential drug targets in animal models of amyotrophic lateral sclerosis (ALS) (Scott et al., 2008), women’s health and cardiovascular disease (Prinz et al., 2011), cancer (Begley and Ellis, 2012; Prinz et al., 2011) or spinal cord injury (Steward et al., 2012), the authors of follow-up studies were unable to replicate most data published by others, including those published in high-profile journals. This failure to replicate has been attributed to a number of factors, including chance observations due to small sample size, concerns about study quality and transparency in reporting, lack of robustness or generalization of the original findings, and publication bias [see Brunner et al. for a discussion of statistical ramifications associated with publication bias (Brunner et al., 2012)]. These and other alarming data have prompted a number of disease-focused groups, including the NINDS, to publish guidelines and policy statements for improving the quality and transparency of preclinical animal research (see Box I).

Moreover, the NINDS considers independent replication to be an asset in evaluating the readiness of a candidate therapy for NINDS-supported translational programs or clinical trials (NINDS Notice NOT-NS-023; http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf). Ideally, replication should be conducted by an independent laboratory (or by a contract research
Box 1. A call for increased rigor and transparent reporting in preclinical research

Given the remarkably low success rate in translating preclinical research into clinical success, a number of disease-focused groups, including the NINDS, are recognizing the urgent need to raise the standards of preclinical studies to encompass the rigor and transparency that is already expected of human clinical trials (see CONSORT 2010 statement (Schulz et al., 2010)). Recently published guidelines and policy statements from these groups outline some of the principles and standards of good study design and reporting when conducting preclinical trials of candidate therapeutics – e.g. allocation concealment, blinded assessment of outcome, random allocation of subjects to experimental groups and other methods designed to minimize bias and Type 1 (false positive) errors (ARRIVE Guidelines, 2010 (http://www.nc3rs.org.uk/page.asp?id=1357); STAIR preclinical recommendations (Fisher et al., 2009); NINDS Notice NOT-NS-11-023 (http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf); Alzheimer's disease preclinical research guidelines (Shineman et al., 2011)). In addition to publishing guidance (NINDS Notice NOT-NS-11-023), the NINDS has incorporated these guidelines into enhanced review considerations for NINDS-supported translational (e.g. http://grants.nih.gov/grants/guide/pa-files/PAR-11-294.html) and clinical (e.g. http://grants.nih.gov/grants/guide/pa-files/PAR-11-343.html) research programs. Moreover, the NINDS convened a 2-day workshop (20-21 June 2012 in Washington, DC) entitled ‘Optimizing the Predictive Value of Preclinical Research’ to identify the key causes of deficiencies in preclinical studies and provide recommendations for addressing them; the outcomes of this workshop have been reported (Landis et al., 2012) and the workshop agenda can be accessed online (http://www.ninds.nih.gov/funding/areas/channels_synapses_and_circuits/rig or_and_transparency/index.htm). By raising standards and awareness, these initiatives strive to increase the reliability, reproducibility and predictive value of preclinical research, and ultimately to improve the likelihood of success on clinical translation.

Disease Models & Mechanisms ● DMM

Translational research in Rett syndrome

Summary of knowledge gaps and workshop recommendations

Despite considerable progress in describing the pathophysiological consequences of Mecp2 mutations in mice, crucial gaps remain that must be addressed in order to establish optimal models for preclinical evaluation of potential RTT therapeutics. The workshop participants concluded that addressing the following research goals in RTT would substantially improve the potential for translating promising preclinical findings into clinical success; these include the need to:

- obtain a thorough understanding of the effects of genetic background on the phenotypic consequences of different Mecp2 alleles
- generate and characterize Mecp2 alleles that model all of the most common human RTT mutations
- develop a detailed characterization of female Het carrying different Mecp2 alleles (cf. Samaco et al., 2012), including how phenotypes appear and evolve across the lifespan; such assessments are crucial given the progressive nature of RTT and the fact that symptoms develop (and change) over the span of years or decades in humans
- better characterize behavioral phenotypes in Mecp2 mutants, particularly those that are relevant to RTT clinical features (e.g. measures of cognitive function, social interactions, communication, anxiety); develop behavioral assays that are not confounded by motor defects
- characterize in more detail the phenotypes that represent significant challenges for medical management of individuals with RTT, including seizures, and gastrointestinal, cardiorespiratory, sensory and sleep-wake problems
- develop more sensitive, reliable and relevant preclinical outcome measures that align with clinical end points and can better predict success in RTT clinical trials; in some cases, this will require a better understanding of whether RTT mouse phenotypes, despite having apparent face validity, share common underlying circuits and neuropathological mechanisms with respective clinical phenotypes
- develop in vivo measures of target engagement (e.g. molecular biomarkers of desired ligand-receptor activity or downstream signaling responses), and biomarkers of disease progression and amelioration that, ideally, are non-invasive, to facilitate translation to clinical studies
- obtain a better understanding of the therapeutic window in RTT – i.e. the optimal timing and duration of treatment in RTT models.

There was agreement at the workshop that outcome measures for preclinical testing in RTT models should involve phenotypes that are robust and reproducible, and ideally possess good construct and face validity with corresponding RTT disease phenotypes in humans. Although mouse phenotypes have been identified that meet these reproducible phenotypes that can serve as a foundation for preclinical testing. Similarly, the therapeutic efficacy of several drugs now used for the treatment of epilepsy, including lacosamide and retigabin, was first established in animal studies organized under the auspices of the Anticonvulsant Screening Program at NINDS (Choi et al., 1996; Rostock et al., 1996; Stöhr et al., 2007). This program incorporates a battery of standardized, rigorously implemented in vivo rodent assays to screen candidate anticonvulsant compounds for epilepsy. Importantly, these examples provide clear evidence that preclinical models have the potential to predict clinical efficacy in humans, provided that the models are robust, studies are well designed and that preclinical outcome measures are relevant to the desired clinical end points.

2013 DMM
criteria (Table 2; supplementary material Tables S3-S7), some phenotypic domains, particularly those involving complex behaviors (e.g. social behaviors and anxiety), still lack consistent, reliable readouts in current RTT models. This prompts an important question: will amelioration of less complex behavioral phenotypes, or even cellular and molecular defects, in preclinical trials be sufficient to predict success in clinical trials involving more complex behavioral end points? This is a possibility, as supported by the recent Phase 2 clinical trial mentioned above involving Fragile X syndrome patients (Berry-Kravis et al., 2012). This study reported improvements in social functioning following treatment with arbaclofen, even though the preclinical data supporting this trial consisted largely of positive drug effects on synaptic and
d

Table 3. Best-practice guidelines for preclinical animal trials testing therapeutic compounds in RTT

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study design</strong></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>Because many studies are frequently underpowered, sample size calculations, based on expected variance and effect size, should be performed before the animal trial starts (i.e. mice should not be added during the trial, once data analysis has begun, in an attempt to increase statistical significance). For most RTT trials, sample sizes should include at least 12-16 mice per group.</td>
</tr>
<tr>
<td>Dose-response data</td>
<td>In addition to the optimal dose, the minimally effective and maximally tolerated dose should be determined. Early evidence that the drug reaches the tissue target (including drug concentration at the target) and preliminary safety and tolerability data can help to support continued development of the candidate therapeutic (e.g. more comprehensive preclinical pharmacokinetics/pharmacodynamics studies).</td>
</tr>
<tr>
<td>Timing and duration of treatment</td>
<td>Treatment onset (e.g. in presymptomatic vs symptomatic mice) and duration (acute vs chronic) should correspond as closely as possible to the expected clinical trial indication. For example, in early Phase 2 RTT clinical trials, the intervention will probably start in individuals who are already symptomatic and will involve chronic intervention. In addition, data on the therapeutic window (e.g. whether the treatment is more effective when initiated in presymptomatic mice) and acute vs chronic treatment effects can be informative in interpreting clinical efficacy data and/or planning later-phase clinical trials.</td>
</tr>
<tr>
<td>Outcome measures</td>
<td>Assessing multiple outcome measures can be informative, especially in exploratory studies. However, for the preclinical animal trial, primary and secondary outcome measures and/or end points should be established in advance, and incorporated into the statistical design of the study. The selection of outcome measure(s) should be as relevant to the expected clinical measure(s) as possible.</td>
</tr>
<tr>
<td>Biomarkers</td>
<td>If available, biomarkers that measure target engagement or activity, and disease progression or amelioration, can be informative in animal studies, especially if the biomarker is translatable to human studies.</td>
</tr>
<tr>
<td><strong>Statistics and minimizing bias</strong></td>
<td></td>
</tr>
<tr>
<td>Statistical design</td>
<td>Statistical methods should be chosen before the study has begun. Once the outcome variability has been characterized, it is a good idea to consult with a statistician on the design of the trial.</td>
</tr>
<tr>
<td>Randomization and balancing</td>
<td>The study design should include how mice will be randomly allocated to treatment groups (e.g. computer-generated randomization schedules). Picking animals ‘at random’ from a cage is likely to introduce bias based on subtle (or not so subtle) characteristics of the animal. Treatment groups should be balanced by age and sex of mice, if relevant, and to prevent over-representation of sibs in any experimental cohort.</td>
</tr>
<tr>
<td>Blinding</td>
<td>Wherever possible, individuals conducting the experiments and those analyzing the results should be blinded to the experimental group (i.e. allocation concealment and blinded assessment of outcome, respectively).</td>
</tr>
<tr>
<td>Inclusion/exclusion criteria</td>
<td>Determine in advance the criteria by which animals will be included or excluded from the study, or discarded from analysis (e.g. based on phenotype, disease severity, sickness or other factors that could skew the outcomes). Post-hoc exclusion of animals or data after the analysis (before unblinding) should be justified and reported.</td>
</tr>
<tr>
<td><strong>Reporting</strong></td>
<td></td>
</tr>
<tr>
<td>ARRIVE guidelines</td>
<td>We suggest consulting the ARRIVE (Animal Research: Reporting In Vivo Experiments) guidelines to report RTT preclinical animal trial results, and to include details on mouse model and genetic background, animal housing and handling, sample size calculations and statistical methods, methods to minimize bias (randomization, blinding), inclusion/exclusion criteria, animals and data excluded from analysis, and other relevant items on the checklist provided in the ARRIVE guidelines.</td>
</tr>
<tr>
<td>Negative results</td>
<td>Publish all results (positive and negative outcomes) to avoid publication bias.</td>
</tr>
<tr>
<td>Conflict of interest</td>
<td>Report any relationship that could be perceived as a conflict of interest, or that could potentially bias the study outcomes.</td>
</tr>
<tr>
<td><strong>Replication</strong></td>
<td></td>
</tr>
<tr>
<td>Multiple RTT models</td>
<td>To evaluate the robustness and generality of the original findings, we suggest validating promising treatment outcomes in more than one RTT mouse model/strain (e.g. comparing the Mecp2 null and a disease-specific mutation, mice with different genetic backgrounds, etc.). Moreover, validating findings in female heterozygous Mecp2 mice is imperative.</td>
</tr>
<tr>
<td>Independent replication</td>
<td>Ideally, an independent laboratory (or a contract research organization) with no financial, scientific or other vested interest in the original study should conduct the replication study. The replication study can attempt to duplicate exactly the original experimental design or modify the procedures (e.g. using a different RTT mouse model/strain) to evaluate the robustness and generality of the findings. All results (positive and negative) should be reported.</td>
</tr>
</tbody>
</table>
neuronal phenotypes in Fmr1 mutant mice (Henderson et al., 2012). Hopefully, as the RTT field moves into clinical testing of therapeutic leads, we can validate and refine the predictive potential of the models by comparing outcomes in mice with results obtained from human clinical trials.

To assess the reproducibility and robustness of promising preclinical interventions, the workshop participants recommended independent replication of the original findings and validation in more than one RTT mouse model or strain prior to launching later-stage translational or clinical projects. In addition, if the original studies were performed in male Nulls, it will be vital to replicate the findings in female Hets whenever feasible, because these models have better construct and face validity for RTT, as discussed above. Adopting standardized assays, tests and outcome measures in animal studies has also been debated in the field (Paylor, 2009; van der Staay and Steckler, 2002). At the workshop, there was agreement that standardized testing procedures can be valuable, especially for phenotypic characterizations (to allow head-to-head comparisons of a phenotype across mouse models or strains, for example); however, overly rigid adherence to ‘accepted’ standards could ultimately inhibit the development or adoption of more sensitive and reliable measures. The workshop participants recommended adopting standardized testing procedures and data analysis protocols whenever possible so that results can be meaningfully compared across different interventions and among different laboratories; modifications and/or improvements to standardized protocols should be justified.

The workshop participants recognized that research consortia, perhaps organized around specific themes (e.g. cognitive function, respiratory control, etc.) could be a valuable tool for generating standardized testing protocols, developing or improving preclinical outcome measures and cross-validating preclinical trial results using best practices for study design (described above). Support for such activities could potentially fall within the scope of the current NIH funding initiative PAR-11-038 (http://grants.nih.gov/grants/guide/pa-files/PAR-11-038.html). The participants also felt that consideration should be given to evaluating the potential value and cost effectiveness of centralized core facilities for preclinical trials, and for validating results from other laboratories.

Finally, we need to bring the rigor and quality of study design that is expected of human clinical trials [see CONSORT 2010 statement (Schulz et al., 2010)] to animal trials that test candidate therapeutics. This will require a commitment among RTT researchers to fully disclose all aspects of study design in their publications, including (but not limited to) allocation concealment, blinded assessment of outcome, random allocation of subjects to experimental groups, experimental group sizes, breeding strategies and statistical methods. Table 3 outlines best-practice guidelines for animal trials in RTT that were derived from the workshop recommendations, and adapted from the STAIR (Fisher et al., 2009), ARRIVE (http://www.nc3rs.org.uk/page.asp?id=1357) and NINDS [NINDS Notice NOT-NS-11-023 (http://www.ninds.nih.gov/funding/transparency_in_reporting_guidance.pdf)] guidelines.

Conclusions

The compelling need for effective treatments for RTT, coupled with the availability of good mouse models, is fuelling interest in translational studies aimed at identifying potential new therapeutics. However, given the substantial financial costs and high expectations of clinical trials, it is incumbent upon the RTT research community to rigorously validate models, outcome measures and study designs (see Table 3) that will generate robust and reproducible preclinical findings with clear relevance to the human disease. The NIH, as well as private Rett syndrome funding organizations (IRSF and RSRT), recognize the added demands that rigorous high-quality animal trials place on investigators, especially in requesting independent replication of promising therapeutic leads before moving forward. Current NIH research funding initiatives target these priorities for both animal (http://grants.nih.gov/grants/guide/pa-files/PAR-11-038.html) and clinical (http://grants.nih.gov/grants/guide/pa-files/PAR-11-045.html) research in RTT.

It is also imperative that the complexity of RTT pathophysiology is taken fully into account in the design of preclinical studies. This includes the recognition that loss of MeCP2 function can have direct effects on neuronal function (e.g. synaptic strength) and development (e.g. neuronal cell growth), as well as indirect and cumulative effects on long-term maturation, leading to distinct phenotypic consequences and possibly requiring different treatment strategies. Similarly, it will be important to determine in RTT individuals which specific features might be reversible with pharmacological treatment alone (and at what developmental age) and which might require more complex interventions (e.g. pharmacotherapy combined with behavioral, cognitive and/or speech therapies). Understanding this complexity will be essential in selecting therapeutic end points and to understanding how best to reach them.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge contributions from Monica Coenraads (RSRT), Janice Ascano (IRSF), Stephen Bajardi (IRSF), John McCall (PharMac LLC), Melissa Parish (NICH) and MaryLou Oster-Granite (NICHID) in organizing the workshop. We thank all workshop speakers and participants (see supplementary material Table S2) for productive discussions, many of which are represented in this paper. We also thank Shai Silberberg (NINDS) and Naomi Kleitman (NINDS) for helpful comments on the manuscript, Nino Ramirez for input on respiratory phenotypes in RTT, and Elsa Perez for compiling the bibliography. Special thanks go to the editorial staff at Disease Models & Mechanisms for attending the workshop and providing invaluable support for this article. This article represents the views of the authors and not the NIH.

COMPETING INTERESTS

The authors declare that they do not have any financial or competing interests.

FUNDING

Support for the September 2011 workshop on which this paper reports (‘Setting Priorities for Therapy Development in Rett Syndrome’) was provided by the National Institute of Neurological Disorders and Stroke (NINDS), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD), the International Rett Syndrome Research Foundation (IRSF) and the Rett Syndrome Research Trust (RSRT). Grant support includes NINDS (NS-057398 to D.M.K. and NS-065027 to L.P.-M.), Rett Syndrome Research Trust (to M.J.J.), the International Rett Syndrome Foundation (ANGEL award 2608 to M.J.J. and ANGEL award 2583 to D.M.K.) and Canadian Institutes of Health (MOF-106481 to J.H.E.).

SUPPLEMENTARY MATERIAL

Supplementary material for this article is available at http://dmm.biologists.org/lookup/doi/10.1242/dmm.011007/-/DC1

REFERENCES


Brunner, D., Balcı, F. and Ludvig, E. A. (2001). Deficiency of methyl-
Brendel, C., Belakhov, V., Werner, H., Wegener, E., Gärtner, J., Nudelman, I.,
Baasov, T. and Huppke, P. (2010). The disease
determination of rett syndrome.
Baptista, P. M., Mercadante, M. T., Macedo, E. C. and Schwartzman, J. S.
Asaka, Y ., Jugloff, D. G., Zhang, L., Eubanks, J. H. and Fitzsimonds, R. M.
Disease Models & Mechanisms


Tyrosine hydroxylase deficit in the hippocampal region of mice. Brain 129, 887-898.


