Analysis of the phenotypes in the Rett Networked Database.

Elisa Frullanti[1], Filomena T. Papa[1], Elisa Grillo[1], Angus Clarke[2], Bruria Ben Zeev[3], Mercedes Pineda[4], Nadia Bahi-Buisson[5-7], Thierry Bienvenu[6-8], Judith Armstrong[9], Ana Roche Martinez[4], Francesca Mari[1,10], Andreea Nissenkorn[3], Caterina Lo Rizzo[10], Edvige Veneselli[11], Silvia Russo[12], Aglaia Vignoli[13], Giorgio Pini[14], Milena Djuric[15], Anne-Marie Bisgaard[16], Vlatka Mejaski Bosnjak[17], Joussef Hayek[18], Rajni Khajuria[19], Barbara Montomoli[20], Francesca Cogliati[12], Maria Pintaudi[21], Kinga Hadzsiev[22], Dana Craiu[23], Victoria Voinova[24], Alessandra Renieri[1,10*], Laurent Villard[25-26*].

1. Medical Genetics, University of Siena, Italy.
2. Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, Wales, UK.
3. Pediatric Neurology Unit and Israeli Rett Clinic, Safra Children Hospital, Chaim Sheba Medical Center, Tel Ha Shomer, Israel; Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel.
4. Neurologia Fundació Sant Joan de Deu, Barcelona, Spain.
5. Pediatric Neurology, Necker Enfants Malades Hospital, Université Paris Descartes, Paris, France.
6. INSERM, U1016, Paris, France.
7. Institute Cochin, Université Paris Descartes, CNRS, UMR8104, Paris, France.
8. Laboratoire de Biochimie et Génétique Moléculaire, Hôpital Cochin, Assistance Publique, Hôpitaux de Paris, Paris, France.
9. Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain; Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain; CIBER-ER (Biomedical Network Research Center for Rare Diseases), Instituto de Salud Carlos III, Madrid, Spain.
10. Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy.
11. Child Neuropsychiatry Unit, Istituto Giannina Gaslini, DINOGMI, University of Genova, Genova, Italy.
12. Istituto Auxologico Italiano, Laboratorio Genetica Molecolare, Milan, Italy.
13. Epilepsy Center, Childhood and Adolescence Neurology and Psychiatry, ASST Santi Paolo Carlo, Department of Health Sciences, University of Milan, Italy
14. Ospedale della Versilia, Viareggio, Italy.
15. Neurologic Department, Mother and Child Health Care Institute of Serbia "Dr Vukan Cupic", University of Belgrade, Belgrade, Serbia.

16. Centre for Rett Syndrome, Department of Paediatrics and Adolescent Medicine, The Kennedy Center, Copenhagen University Hospital Rigshospitalet, Denmark.
17. Department of Neuropediatrics, Zagreb Children's Hospital, University of Zagreb, Zagreb, Croatia.
18. Department of Molecular and Developmental Medicine, University of Siena, Siena, Italy.
19. Genetics Unit, Department of Pediatrics, All India Institute of Medical Sciences, New Delhi, India.
20. Child Neuropsychiatry Unit, Azienda Ospedaliera Universitaria Senese, Siena, Italy.
21. DINOGMI, University of Genova, Genova, Italy.
22. Department of Medical Genetics, and Szentagothai Research Center, University of Pécs, Pécs, Hungary.
23. Carol Davila University of Medicine, Pediatric Neurology Clinic, Al Obregia Hospital, Bucharest, Romania.
24. Department of Clinical Genetics, Institute of Pediatrics and Pediatric Surgery, Ministry of Health, Moscow, Russia.
25. Aix Marseille Université, GMGF, 13385, Marseille, France.
26. Inserm, UMR_S 910, 13385, Marseille, France.

* Co-corresponding authors

Laurent Villard
Inserm UMR_S 910 - Faculté de Médecine Timone
27 bd Jean Moulin - 13385 Marseille - France
Phone: +33 (0)491324903
Fax: +33 (0)491804319
e-mail: laurent.villard@univ-amu.fr

Alessandra Renieri
Medical Genetics
University of Siena, Policlinico S Maria alle Scotte
Viale Bracci, 2 - 53100 - Siena - Italy
Phone: +39 0577 233303
FAX: +39 0577 233325
e-mail: renieri@unisi.it

ABSTRACT

BACKGROUND: Rett spectrum disorder is a progressive neurological disease and the most common genetic cause of mental retardation in females. *MECP2*, *CDKL5*, and *FOXG1* have been found to be the causative genes. Each gene and different mutations within each gene contribute to variability in clinical presentation and several groups worldwide performed genotype-phenotype correlation studies using cohorts of patients with classic and atypical forms of Rett spectrum disorder.

OBJECTIVE and METHODS: The Rett Networked Database is a unified registry of clinical and molecular data of Rett patients and it is currently one of the largest Rett registry worldwide with more than 1900 records provided by Rett expert clinicians from 13 countries.

RESULTS: Collected data revealed that the majority of *MECP2* mutated patients present with the classic form, the majority of *CDKL5* mutated patients with the early onset seizure variant and the majority of *FOXG1* mutated patients with the congenital form. A computation of severity scores further revealed significant differences between groups of patients and correlation with mutation types.

CONCLUSIONS: The highly detailed phenotypic information contained in the Rett Networked Database allow to group patients presenting specific clinical and genetic characteristics for studies by the Rett community and beyond. These data will also serve for the development of clinical trials involving homogeneous groups of patients.

Keywords: patient registry, Rett syndrome, genotype-phenotype correlation.

INTRODUCTION

Rett syndrome (RTT, OMIM 312750) is a severe X-linked neurodevelopmental disorder that affects predominantly females with an incidence of approximately 1 in 10,000 female births. [1,2,3]. Classic RTT is infrequently observed in males because a deleterious mutation in the only copy of *MECP2* typically results in severe neonatal encephalopathy and early lethality [4]. In the classic form, girls with RTT typically exhibit a relatively normal period of development for the first 6-18 months of life followed by a regression phase where patients lose acquired language and motor skill and exhibit intellectual disability and hand stereotypies. The hand stereotypies are typical in RTT and appear to be continuous, located predominantly over the anterior body midline [5]. Beyond the classic form of RTT, a number of atypical forms with different degrees of severity have been described: the Zappella variant (formerly known as preserved speech variant) [6,7], the infantile seizure onset type [8], the congenital form [9] and the 'forme fruste'[10]. The Rett Networked Database (RND) is a registry of clinical and molecular data for patients affected by RTT [11]. Although it was initially targeting the European population of patients with RTT, it is now open to countries outside of Europe. RND is different from existing repositories for RTT clinical and molecular data [2,12,13]. First, RND records are updated by clinicians with experience in RTT. This is important to exclude potential bias existing when clinical data are gathered using questionnaires sent out to families by mail. Second, RND is among the largest RTT registries worldwide with more than 1900 patients on file. Third, RND is an open access initiative and data can be retrieved directly through a web-based search engine by interested professionals upon the submission of a research proposal to the Scientific Review Board [11] Public has access to general information and to content

description while the individual patient file can be granted only upon registration of physicians and clinical researchers in charge of specific patients. Here, we describe the first 1925 records contained in the registry and discuss the content of RND on the basis of the published guidelines for RTT clinical diagnosis. We analyzed the phenotype of patients with *MECP2*, *CDKL5* and *FOXG1* mutation to better understand the typical and atypical forms of RTT and provided information of RTT cohorts for the development of clinical trials.

MATERIALS AND METHODS

RND data

RND contains clinical files for 1958 patients affected by classic or atypical RTT (numbers are given as of March 1st, 2017). Clinical data originates from Croatia (29 patients), Denmark (64 patients), France (252 patients), Hungary (82 patients), India (3 patients), Israel (93 patients), Italy (605 patients), Romania (15 patients), Serbia (50 patients), Spain (398 patients), United Kingdom (255 patients), USA (96 patients) and Russia (16 patients).

Statistical Analysis

Differences in clinical characteristics between groups of patients were tested by Fisher's exact test or by chi-squared analysis when the normal approximation was appropriate.

RESULTS

Overview of RND data

One thousand nine hundred fifty-eight patients with classic or atypical RTT are actually contained in RND. Among these, 1508 patients (77,0%) carry a mutation in a disease-

causing gene while 450 patients (23%) do not show mutations in any of the gene linked to RTT syndrome. As expected, *MECP2* is mutated in the majority of patients with 1404 patients carrying a *MECP2* mutation (93,1%). 78 patients carrying a mutation in *CDKL5* (5,2%) and 26 patients carrying a mutation in *FOXG1* (1,7%), much less frequent causes of RTT syndrome. All cases are sporadic except for 2 pairs of sisters and 5 pairs of monozygotic twins affected by RTT and carrying a *MECP2* mutation.

Patients carrying a mutation in *MECP2*

RND contains clinical files for 1404 patients carrying a mutation in the *MECP2* gene. The mutation type and clinical variant type are available for 1141 patients (1131 females and 10 males). All mutation types are present in this population with early truncating and missense mutations being the most frequent mutations despite significant difference between classical and atypical forms (Table 1). A small percentage of patients (0,5%) has a mutation whose functional significance is unknown. This category includes synonymous mutations and intronic variants outside donor/acceptor splice sites potentially affecting the splicing of *MECP2* and for which not enough functional data are available in order to support pathogenicity. These mutations are absent from the databases of variants identified in control individuals. Nine hundred twenty-one patients have the classic form of RTT (80,7%), 117 patients have the Zappella variant (10,3%), 46 patients have the congenital variant of RTT (4%), seven patients have the early onset seizure variant (0,7%). Twenty-one patients (1,8%) do not fall into any of these categories because of the lack of one or several diagnostic criteria necessary for classical or atypical form classification. In addition, 19 patients (1,7%) are too young to be clinically defined as affected by the Zappella variant or the classical form. Criteria for the clinical diagnosis of RTT were last revised in RTT Diagnostic Criteria 2010 [14] in

order to include a regression period, partial or complete loss of acquired purposeful hand skills, stereotypic hand movements, partial or complete loss of acquired spoken language and gait abnormalities. We mined the RND data in order to investigate their compliance with the revised diagnostic criteria. Our analysis showed that, among the patients carrying a mutation in *MECP2*, regression occurred in 94,9% of patients, 90,6% lost or never acquired purposeful hand skills, 82,1% have stereotypic hand movements, 70,5% lost most or all spoken language and 74% have gait dyspraxia (Table 2). Notably, hand stereotypies, although considered an invariant clinical sign of classical RTT, are absent in 17,9% of *MECP2* mutated patients included in the RND dataset. On the other hand, the intense eye pointing phenotype of RTT patients, is present in 88% of *MECP2* positive cases (Table 2), although not included in the necessary criteria. RND data were further interrogated to define the most frequent clinical signs, i.e. clinical signs present in more than a 50% of *MECP2* mutation carriers (Table 3). This analysis revealed that, in addition to the necessary criteria for RTT diagnosis, a normal head circumference at birth (93,5%) and a deficient sphincter control (88,7%) are main clinical signs in *MECP2*-mutated patients. Interestingly, although 85,2% of the *MECP2* positive patients have feeding difficulties (sometimes requiring gastrostomy), only half of them (53,6%) have gastrointestinal disturbances. This would suggest that part of the feeding difficulties arise from abnormal muscle tone and oropharyngeal dysfunction [15]. Even though breathing mechanisms in RTT preclinical models have been heavily investigated, breathing dysfunction "only" affects half (50%) of the patients carrying a mutation in *MECP2*. Interestingly, although loss of acquired speech is included among RTT diagnostic criteria, RND data show that the majority of *MECP2* positive cases have never spoken (57,1%), as reported in Table 3. Among *MECP2*-mutated patients, it was possible to

compute the total score for 584 of the 1141 individuals for which the mutation status is available. This group included subjects over 5 years of age, and the most recent examination was used to compute the score. A cumulative distribution plots of patients positive for the *MECP2* mutations is illustrated in Figure 1A and it is grouped on the basis of the four common point mutations (Arg106Trp, Arg133Cys, Thr158Met, Arg306Cys), the four early truncating mutations (Arg168*, Arg255*, Arg270*, Arg294*), late truncating mutations (LTM), large deletions and all other mutations. Although there is wide variability in clinical severity, there is a clear effect of specific common *MECP2* point mutations on median clinical severity (Figure 1A). We indeed reported a case of 2 pairs of sisters carrying the same *MECP2* mutation but with discordant clinical signs. One individual from each pair could not speak or walk, and had a profound intellectual deficit (classical RTT), while the other individual could speak and walk, and had a moderate intellectual disability (Zappella variant) [16]. The 5 monozygotic twin pairs reported in RND, are much more concordant than the sister pairs. However, it is interesting to note that only 2 out of 5 twin pairs have an identical clinical score, indicating that at least at this level of investigation they are phenotypically identical. The 3 other twin pairs differ in specific fields such as epilepsy and weight (twin pair 1), level of speech and level of phrases (twin pair 2), or height, age of regression and voluntary hand use (twin pair 3).

Patients carrying a mutation in *CDKL5*

The phenotype of the patients carrying a mutation in *CDKL5* and classified as having atypical RTT is much less documented than the classical RTT phenotype caused by *MECP2* mutations. A report in 2013 described 86 patients with a mutation in *CDKL5* with data originating from family questionnaires [17]. RND contains 78 records for

CDKL5 mutation positive cases (76 females and 2 males). Similar to other records in RND, the clinical and molecular information were provided by experienced clinicians through direct patients' evaluation. The cumulative distribution of the patients positive for *CDKL5* mutations, clustered on the basis of early truncating mutations, late truncating mutations, large deletions, and missense mutations, is illustrated in Figure 1B. Forty-eight patients have the early onset seizure variant of RTT (85,7%), 1 patient has the Zappella variant (1,8%) and 1 patient has the congenital variant of RTT (1,8%). Mutation type is available for 65 patients. The most frequent mutations are truncating mutations (56,9%) followed by missense mutations (27,7%). In our cohort, the majority of patients have a normal head circumference at birth (94,9%), a deficient sphincter control (95,7%), feeding difficulties (97,4%), IQ < 40 (93,2%), presence of hand stereotypies (89,2%) and had never spoken (81,0%) (Table 3). Furthermore, epilepsy before 5 years of age is statistically significant among groups of patients (*p-value* <0.0001 *MECP2* vs. *CDKL5*), since it is present in 100% of *CDKL5* cases. The epilepsy started before 1 year of age in 95,7% of *CDKL5* mutated patients versus 3,9% of *MECP2* mutated patients and 37,5% of *FOXG1* mutated patients. Epilepsy is not controlled by therapy in 74,6% of *CDKL5* cases versus 21,9% of *MECP2* patients and 58,8% of *FOXG*1 mutated patients (Table 3). Involuntary movements, mood disturbances and breathing dysfunction are statistically less frequent in *CDKL5* cases rather than in *MECP2* or *FOXG1* patients. Others features such as normal head circumference at birth, deficient sphincter controls, hand stereotypies, feeding difficulties, height and weight below the 25^{th} percentile and troubled night time sleeping are very similar among the three groups of patients carrying a *MECP2*, *CDKL5* or *FOXG1* mutation (Table 3). As for patients with *MECP2* mutations, it was possible to compute the total score for 37 out

of 65 patients for which *CDKL5* mutation status was available. This group included subjects with epilepsy onset occurring in early infancy for whom the most recent examination was used to compute the score.

Patients carrying a mutation in *FOXG1*

The phenotype of the patients carrying a mutation in *FOXG1* and classified as having atypical RTT is even less documented than the phenotype caused by *CDKL5* mutations. RND contains 27 records for *FOXG1* mutation positive cases (19 females and 8 males). The full score is available for 26 patients. Twenty-three patients have the congenital form of RTT (88,5%), 1 patient has the classical form (3,8%), two do not fall into any of these categories (7,7%) because one or several necessary criteria for the diagnosis of classical or atypical forms were not present at the last examination (included in RND as "Rett-like"). The cumulative distribution of the patients positive for *FOXG1* mutations was obtained on the basis of early truncating mutations, late truncating mutations, gene deletions and missense mutations and it is illustrated in Figure 1C. In our cohort, all patients carrying a *FOXG1* mutation had IQ < 40, microcephaly and no speech at examination (Table 3). Similarly, there is a significant overlap between the phenotype of patients with *FOXG1* mutation and the phenotype of patients with *MECP2* or *CDKL5* mutation. Eye pointing capability is reduced from 88% in *MECP2* to 42% in *CDKL5* and 13,3% in *FOXG1* patients. The percentage of patients having epilepsy before 5 years of age is 65,5%, 100% and 87,5% in the *MECP2, CDKL5* and *FOXG1* group of patients, respectively. However, only *MECP2* versus *CDKL5* and *FOXG1* versus *CDKL5* comparisons were statistically significant. The percentage of patients that had a period of regression, IQ < 40, no speech at examination, scoliosis, breathing dysfunction and epilepsy not controlled by therapy is closer to the percentage observed for *CDKL5*

mutated patients rather than to the percentage observed for *MECP2* patients. The major difference in the *FOXG1* group is the higher percentage of patients that never learned to walk or speak or sit. About 92,3% of patients carrying a *FOXG1* mutation has never spoken compared to the 57,1% of *MECP2* and 81,0% of *CDKL5* mutated patients. Moreover, *FOXG1* mutated patients have never learned to sit (78,3%) and walk (91,3%) compared to *MECP2* (12,9% and 31,7%) and *CDKL5* (29,3% and 61,5%) respectively. The overall cumulative distribution plot of patients carrying a mutation in the *MECP2*, *CDKL5* or *FOXG1* genes is illustrated in figure 2.

Patients without identified mutation

Two hundred sixty-three patients (258 females and 5 males) have no identified mutation. Among these, 247 patients (242 females and 5 males) are negative after mutation screening of genes known to cause RTT phenotypes, 6 cases were not screened and the information is lacking for the remaining 10 patients. Although 120 patients are described as having classical RTT (46,7%) caution is needed in the consideration of such numbers. Indeed, although a majority of patients (89,2%) had a normal head circumference at birth and 88.6% have gone through a period of regression, only 46,8% of patients without mutation had hand stereotypies and 56,7% had never spoken. A third of these without a genetic aetiology (35%) had moderate to mild intellectual disability, when most *MECP2* or *CDKL5* positive cases had an IQ below 40. The most frequent clinical characteristics of patients tested negative for *MECP2*, *CDKL5* or *FOXG1* mutation are presented in Table 4. Obviously, this is a most heterogeneous group and efforts are currently being done to add clinical data to the corresponding files and to increase the molecular screening output.

DISCUSSION

Globally, the majority of RND patients do fulfill necessary criteria for the diagnosis of RTT with >70% of patients complying with the revised criteria [14]. A survey of the Rett syndrome Rare Disease Research Center (RDCRC) encompassing a population of 819 patients diagnosed as having RTT and carrying a mutation in the *MECP2* gene was published [18]. The mutation types in RND and RDCRC cannot be compared because the latter only presents the most frequent mutations and a global percentage of missense versus nonsense mutations is not provided. However, it is interesting to observe that both cohorts have a similar proportion of large deletions (8,7% in RDCRC and 8,5% in RND) in patient with classic RTT. However, a different percentage of classic RTT patients does not carry a mutation in *MECP2* gene (5% in RDCRC and 11,2% in RND). Indeed, the RND population appears to be more clinically heterogeneous than the RDCRC cohort. For example, in RDCRC, 100% of the patients comply with the main criteria of RTT Diagnostic Criteria 2010 [14] necessary for typical or classic RTT (loss of hand use, loss of communication, hand stereotypies and gait abnormalities). In RND, the same criteria are present in 70,5% to 90,6% of classic RTT patients. Surprisingly, RND patients more frequently comply with the revised supportive criteria (all criteria present in > 45% patients) while one third supportive criteria are not found in the majority of RDCRC patients. RND is more "open" than RDCRC and patients can be included even if they do not meet one of the necessary criteria for the diagnosis of "classic" RTT. Based on our experience with more than 1900 patients originating from many different countries, we observed that the diagnosis of "classic RTT", made by the physicians at the moment of clinical files submission to the RND, does not perfectly comply with the RTT Diagnostic Criteria 2010. The RTT Diagnostic Criteria 2010 do not totally account for clinical

observations made by many specialists and these criteria will probably need to be adapted again in a near future. A large cohort (n=86) of patients with a mutation in *CDKL5* was previously described based on the questionnaires collected by the InterRETT database [17]. RND provides the second largest cohort of patients harbouring a mutation in *CDKL5* with 78 cases. Expectedly for the early seizure variant of RTT caused by *CDKL5* mutations, the majority of patients experienced at least one episode of epilepsy (>90% in both cohorts). The proportion of patients with a mutation in *CDKL5* that never learned to walk in the two cohorts is also very similar (67,4% in InterRETT and 61,5% in RND), together with the proportion of patients displaying hand stereotypies (80,3% of females in InterRETT and 89,2% of patients positive for a mutation in *CDKL5* in RND). There is a difference between the two cohorts concerning the speech skills, since 30 out of 76 females with *CDKL5* mutation acquired early speech skills in the InterRETT cohort while only 5/42 females harbouring a *CDKL5* mutation have acquired >10 words at age 5 in RND. In addition, another discrepancy concerns patients that had deceleration of head growth (59,5% in InterRETT and 29% in RND). An earlier study of the North American RTT Database relying on 914 patients with a mutation in *MECP2* was published [13]. However, similar to the Australian database, the data relies on questionnaires sent out to families and even if the questionnaires were analyzed by experienced clinicians, the patients were not all directly examined by the contributors. Available results mainly concern molecular data with the distribution and nature of reported mutations. It does not contain *CDKL5* or *FOXG1* molecular data and does not provide details concerning the major phenotypic traits present in the studied population. In the Percy report [13], 85,5% of patients with *MECP2* mutation have the typical form and 13,4% have the atypical form of RTT. Similarly, the percentage of

typical RTT patients with a *MECP2* mutation in the RND is 80,7% when males and mutations of unknown significance are excluded. The cumulative distribution in Figure 1A illustrates that there is a wide clinical variability within the same *MECP2* mutation. However, the "mildest" mutations are Arg133Cys and late truncating mutations in accordance with previous reports [19]. The missense mutations Arg306, Thr158 and Arg106 (arginine or threonine can be replaced by any amino acid) and the early truncating mutation Arg294* belong to the intermediate severity phenotype. The remaining early truncating mutations (Arg168*, Arg255* and Arg270*) and large deletions are among the "most severe" form of RTT syndrome. It is interesting to note that the plot of each mutation is not always parallel. For example, Thr158Met and Arg294* move more vertical, suggesting that the phenotype of patients who have these mutations is less influenced by other genetic or environmental factors. The cumulative distribution in Figure 1B illustrates that a genotype-phenotype correlation for *CDKL5* may be barely recognized and the lines move together. It is interesting to note that late truncating mutations move less vertical, suggesting that the phenotype of patients who have these mutations is more variable. The cumulative distribution in Figure 1C shows a clear trend toward a less severe phenotype for *FOXG1* late truncating mutations. The cumulative overall distribution in Figure 2 nicely illustrates the progressive severity going from *MECP2* to *CDKL5* and *FOXG1* mutation. *CDKL5* patients lie in the most severe range in comparison to *MECP2* patients with *FOXG1* patients even more shifted than *CDKL5* patients towards a worse clinical phenotype and a very minimum overlap with *MECP2* patients.

In conclusion, the Rett Networked Database is a registry for patients with RTT where clinical data are validated by experienced clinicians upon direct examination of

the affected individuals. One of the unique features of this database is its ability to collect a huge amount of clinical details, being almost 300 the collected clinical items with different levels of completeness, and genetic data [11]. It provides a unique resource in order to perform genotype-phenotype correlations. Overall, observation of RND data highlights clinical characteristics which occur more frequently in patients with a specific mutation (Table 3). For example, presence of regression and gait dyspraxia are statistically more frequent in *MECP2* mutated patients; epilepsy and reduction in eye pointing capability are statistically more frequent in *CDKL5* mutated patients, while the large majority of *FOXG1* patients have never learned to walk, sit and speak. Moreover, we observed that the majority of *MECP2* mutated patients have the classical form of RTT, the majority of *CDKL5* mutated patients have the early onset variant and the majority of *FOXG1* mutated patients have the congenital form, with some exceptions. The overlap between the different phenotypes is presented in Figure 3. RND provides an open structure, available to all interested professionals and a searchable web interface made available for registered users. These characteristics should prove useful to perform additional phenotype-genotype correlations, to better understand the typical and atypical forms of RTT, and to select adequate patient populations for future clinical trials.

ACKNOWLEDGEMENTS

We are grateful to the patients and their families who participated in this study. We thank Rossano Di Bartolomeo and Marco Maria D'Andrea, 3W Net Service, for their work in the construction of the database network. This project was supported by IRSF (microgrants to AR), ERARE EuroRett network, Danish Rett Syndrome Association, Association Française du Syndrome de Rett, Catalan Rett Association, Israel Rett

syndrome Association and Associazione Italiana Rett (AIRETT). The BIRSS is grateful for support from Rett UK. We also thank Rett Syndrome Europe (RSE) for strongly believing in this project.

REFERENCES

[1] WE Kaufmann, JL Stallworth, DB Everman, SA Skinner, Neurobiologically-based treatments in Rett syndrome: opportunities and challenges. Expert Opin Orphan Drugs 4 (2016): 1043-1055.

[2] J Christodoulou, A Grimm, T Maher, B Bennetts, RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. Hum Mutat 21 (2003): 466-472.

[3] A Rett, [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. Wien Med Wochenschr 116 (1966): 723-726.

[4] NC Schanen, TW Kurczynski, D Brunelle, MM Woodcock, LSt Dure, AK Percy, Neonatal encephalopathy in two boys in families with recurrent Rett syndrome. J Child Neurol 13 (1998): 229-231.

[5] S Goldman, T Temudo, Hand stereotypies distinguish Rett syndrome from autism disorder. Mov Disord 27 (2012): 1060-1062.

[6] M Zappella, The Rett girls with preserved speech. Brain Dev 14 (1992): 98-101.

[7] A Renieri, F Mari, MA Mencarelli, E Scala, F Ariani, I Longo, I Meloni, G Cevenini, G Pini, G Hayel, M Zappella, Diagnostic criteria for the Zappella variant of Rett syndrome (the preserved speech variant). Brain Dev 31 (2009): 208-216.

[8] F Hanefeld, The clinical pattern of the Rett syndrome. Brain Dev 7 (1985): 320-325.

[9] S Rolando, Rett syndrome: report of eight cases. Brain Dev 7 (1985): 290-296.

[10] B Hagberg, P Rasmussen, "Forme fruste" of Rett syndrome--a case report. Am J Med Genet Suppl 1 (1986): 175-181.

[11] E Grillo, L Villard, A Clarke, B Ben Zeev, M Pineda, N Bahi-Buisson, A Hryniewiecka-Jaworska, T Bienvenu, J Amstrong, A Roche-Martinez, F Mari, E

Vesenelli, S Russo, A Vignoli, G Pini, M Djuric, AM Bisgaard, V Mejaski Bosnjak, N Polgar, F Cogliati, K Ravn, M Pintaudi, B Melegh, D Craiu, A Djukic, A Renieri, Rett networked database: an integrated clinical and genetic network of Rett syndrome databases. Hum Mutat 33 (2012): 1031-1036.

[12] S Fyfe, A Cream, N de Klerk, J Christodoulou, H Leonard, InterRett and RettBASE: International Rett Syndrome Association databases for Rett syndrome. J Child Neurol 18 (2003): 709-713.

[13] AK Percy, JB Lane, J Childers, S Skinner, F Annese, J Barrish, E Caeg, DG Glaze, P MacLeod, Rett syndrome: North American database. J Child Neurol 22 (2007): 1338-1341.

[14] JL Neul, WE Kaufmann, DG Glaze, J Christodoulou, AJ Clarke, N Bahi-Buisson, H Leonard, ME Bailey, NC Schanen, M Zappella, A Renieri, P Huppke, AK Percy, RettSearch Consortium, Rett syndrome: revised diagnostic criteria and nomenclature. Ann Neurol 68 (2010): 944-950.

[15] KJ Motil, RJ Schultz, K Browning, L Trautwein, DG Glaze, Oropharyngeal dysfunction and gastroesophageal dysmotility are present in girls and women with Rett syndrome. J Pediatr Gastroenterol Nutr 29 (1999): 31-37.

[16] E Grillo, C Lo Rizzo, L Bianciardi, V Bizzarri, M Baldassarri, O Spiga, S Furini, C De Felice, C Signorini, S Leoncini, A Pecorelli, L Ciccoli, MA Mencarelli, J Hayek, I Meloni, F Ariani, F Mari, A Renieri, Revealing the complexity of a monogenic disease: rett syndrome exome sequencing. PLoS One 8 (2013): e56599.

[17] S Fehr, M Wilson, J Downs, S Williams, A Murgia, S Sartori, M Vecchi, G Ho, R Polli, S Psoni, X Bao, N de Klerk, H Leonard, J Christodoulou, The CDKL5

disorder is an independent clinical entity associated with early-onset encephalopathy. Eur J Hum Genet 21 (2013): 266-273.

[18] AK Percy, JL Neul, DG Glaze, KJ Motil, SA Skinner, O Khwaja, HS Lee, JB Lane, JO Barrish, F Annese, L McNair, J Graham, K Barnes, Rett syndrome diagnostic criteria: lessons from the Natural History Study. Ann Neurol 68 (2010): 951-955.

[19] M Zappella, I Meloni, I Longo, G Hayek, A Renieri, Preserved speech variants of the Rett syndrome: molecular and clinical analysis. Am J Med Genet 104 (2001): 14-22.

TABLES

Table 1. *MECP2* patients classified as classical or atypical RTT reported in RND (total number of patients 1101). Cases below age 4 (n=19) are not taken into account here. Twenty-one patients do not fall into any of these categories because one or several necessary criteria for the diagnosis of classical or atypical forms were not present at last examination. In RND, early truncations are the mutations altering the MECP2 protein before amino acid 310. N represents the number of cases for which the corresponding item is present in the patient file, percentage is provided in brackets.

Mutation Type	Classic RTT N (%classical)	Atypical RTT N (%atypical)	Males N	P-value classical vs. atypical RTT	Total (%)
Early truncating	404 (43,9)	36 (22)	3	<0,0001	443 (40,2)
Missense	328 (35,7)	86 (50)	2	0,0002	416 (37,8)
Late truncating	107 (11,6)	40 (23,2)	2	<0,0001	149 (13,5)
Gene deletion	78 (8,5)	6 (3,6)	1	0,0264	85 (7,7)
Gene duplication	-	1 (0,6)	2	0,0199	3 (0,3)
Unknown significance	4 (0,3)	1 (0,6)	-	0,7848	5 (0,5)

Table 2. Compliance of the RND data with the revised diagnostic criteria [14] for patients positive for a mutation in *MECP2*. Peripheral vasomotor disturbances are accounted for in the "Small cold hands and feet" score. The item "Diminished response to pain" is not present in RND data. N represents the number of cases for which the corresponding item is present in the patient file. N+ represents the number of cases positive for the clinical signs, percentage is provided in brackets.

Clinical Sign	**N**	**N+ (%)**
A period of regression	1189	1129 (94,9)
NECESSARY CRITERIA		
Partial or complete loss of acquired purposeful hand skills	961	871 (90,6)
Stereotypic hand movements	1185	973 (82,1)
Partial or complete loss of acquired spoken language	1039	733 (70,5)
Gait abnormalities	700	518 (74,0)
SUPPORTIVE CRITERIA		
Breathing disturbances when awake	1003	502 (50,0)
Bruxism when awake	829	515 (62,1)
Impaired sleep pattern	926	419 (45,2)
Abnormal muscle tone	-	-
Peripheral vasomotor disturbances	-	-
Scoliosis or kyphosis	1189	754 (63,4)

Growth retardation*	1062	608 (57,3)
Small cold hands and feet	1042	642 (61,6)
Inappropriate laughing or screaming spells	560	171 (30,5)
Diminished response to pain	-	-
Eye pointing	958	843 (88,0)

*Growth retardation was considered to be present when weight was below the 25th percentile. When height is considered, 61,6% (n=635) of *MECP2* positive patients are below the 25th percentile.

Table 3. Main clinical characteristics in patients positive for a mutation in *MECP2*, *CDKL5*, and *FOXG1*. Clinical characteristics are listed in descending order of percentage of patients harbouring a mutation in *MECP2*. N represents the number of cases for which the corresponding item is present in the patient file. N+ represents the number of positive for the clinical sign, percentage is provided in brackets. *P-value* of significance is provided for comparison.

Clinical Sign	***MECP2***		***CDKL5***		***FOXG1***		***P-value MECP2* vs. *CDKL5***	***P-value MECP2* vs. *FOXG1***	***P-value CDKL5* vs. *FOXG1***
	N	**N+ (%)**	**N**	**N+ (%)**	**N**	**N+ (%)**			
A period of regression	1189	1129 (94,9)	61	27 (44,3)	25	13 (52)	<0,0001	<0,0001	0,5136
Normal head circumference at birth	900	842 (93,5)	39	37 (94,9)	17	14 (82,4)	0,7421	0,0663	0,1309
Deficient sphincter control	974	864 (88,7)	46	44 (95,7)	23	22 (95,7)	0,1409	0,2952	1
Eye pointing	958	843 (88,0)	50	21 (42)	15	2 (13,30)	<0,0001	<0,0001	0,0417

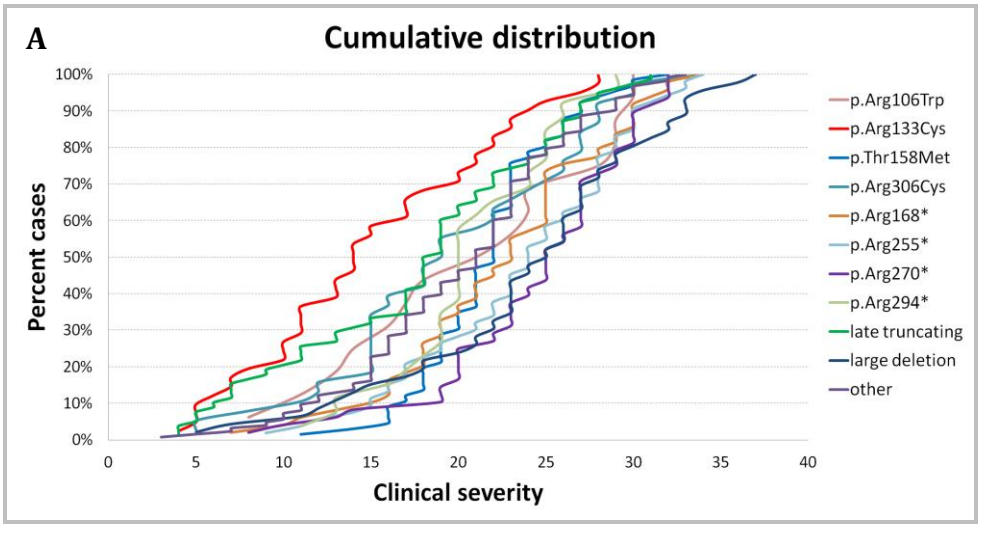
Feeding difficulties	813	693 (85,2)	38	37 (97,4)	15	13 (86,7)	0,0364	0,8772	0,1288
Presence of hand stereotypies	1185	973 (82,1)	65	58 (89,2)	24	23 (95,8)	0,1414	0,0806	0,3338
IQ < 40	916	701 (76,6)	59	55 (93,2)	24	24 (100)	0,0029	0,0069	0,191
Microcephaly or deceleration of head growth	1110	837 (75,4)	69	28 (55,1)	26	26 (100)	<0,0001	0,0037	<0,0001
Gait dyspraxia	700	518 (74,0)	25	13 (52,0)	21	2 (33,3)	0,0146	<0,0001	0,0022
No speech at examination	1039	733 (70,5)	42	37 (88,1)	24	24 (100)	0,0138	0,0016	0,0787
Epilepsy before 5 years of age	689	451 (65,5)	64	64 (100)	16	14 (87,5)	<0,0001	0,0658	0,0042
Scoliosis	1189	754 (63,4)	40	9 (22,5)	22	6 (27,3)	<0,0001	0,0033	0,6746
Bruxism	829	515 (62,1)	49	21 (42,9)	17	11 (64,7)	0,0072	0,0782	0,0034

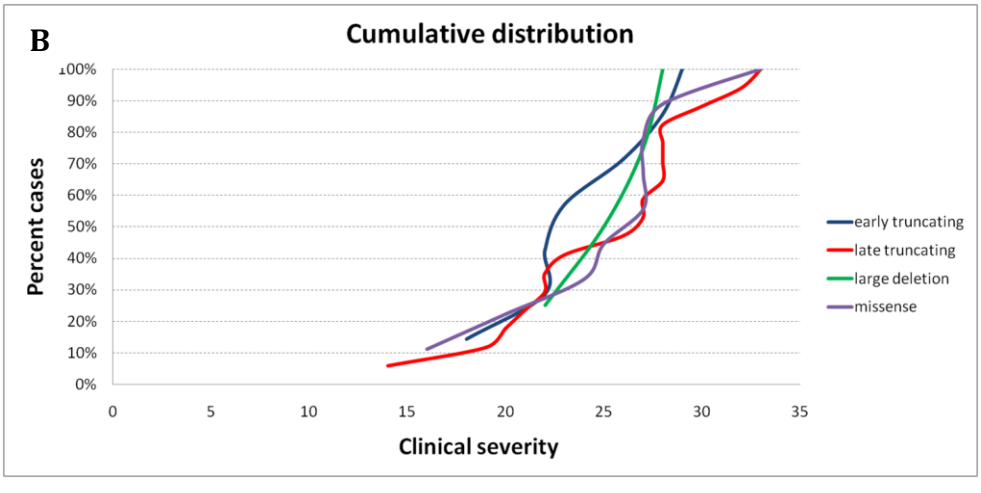
Height below the 25th percentile	1031	635 (61,6)	42	20 (47,6)	25	17 (68)	0,0687	0,5147	0,1047
Cold extremities	1042	642 (61,6)	41	16 (39)	22	10 (45,5)	0,0037	0,1236	0,6212
Weight below the 25th percentile	1062	608 (57,3)	46	22 (47,8)	25	17 (68)	0,2064	0,2825	0,1027
Has never spoken	1039	593 (57,1)	42	34 (81,0)	26	24 (92,3)	0,0021	0,0003	<0,0001
Involuntary movements	781	429 (54,9)	24	4 (16,7)	12	8 (66,7)	0,0002	0,4172	0,0027
Gastrointestinal disturbances	714	383 (53,6)	37	16 (43,2)	22	16 (72,7)	0,2165	0,0768	0,0279
Mood disturbance	787	398 (50,6)	39	11 (28,2)	17	12 (70,6)	0,0064	0,1024	0,003
Breathing dysfunction	1003	502 (50,0)	43	6 (13,9)	25	7 (28)	<0,0001	0,0294	0,1555
Troubled night time sleeping	964	431 (44,7)	62	27 (43,5)	18	11 (61,1)	0,8585	0,1658	0,189
Never learned to walk	1180	374	65	40	23	21	<0,0001	<0,0001	0,0078

		(31,7)		(61,5)		(91,3)			
Epilepsy not controlled by therapy	1018	223 (21,9)	71	53 (74,6)	17	10 (58,8)	<0,0001	0,0003	0,1938
Never learned to sit	1095	141 (12,9)	41	12 (29,3)	23	18 (78,3)	0,0025	<0,0001	0,0002
Epilepsy before 1 years of age	689	27 (3,9)	69	66 (95,7)	16	6 (37,5)	<0,0001	<0,0001	<0,0001

Table 4. Clinical characteristics in patients negative for mutation in *MECP2*, *CDKL5* or *FOXG1*. The table is limited to the clinical characteristics present in more than 50% of these patients. N represents the number of cases for which the corresponding item is present in the patient file. N+ represents the number of positive for that clinical sign, percentage is provided in brackets.

Clinical Sign	***MECP2* negative cases**	
	N	**N+ (%)**
Deficient sphincter control	228	205 (90,0)
Normal head circumference at birth	167	149 (89,2)
A period of regression	245	217 (88,6)
Regression of hand use	229	187 (81,7)
Presence of epilepsy before age 5	216	166 (76,9)
Microcephaly or deceleration of head growth	229	159 (69,4)
Bruxism	234	154 (65,8)
IQ < 40	203	132 (65,0)
Has never spoken	240	136 (56,7)





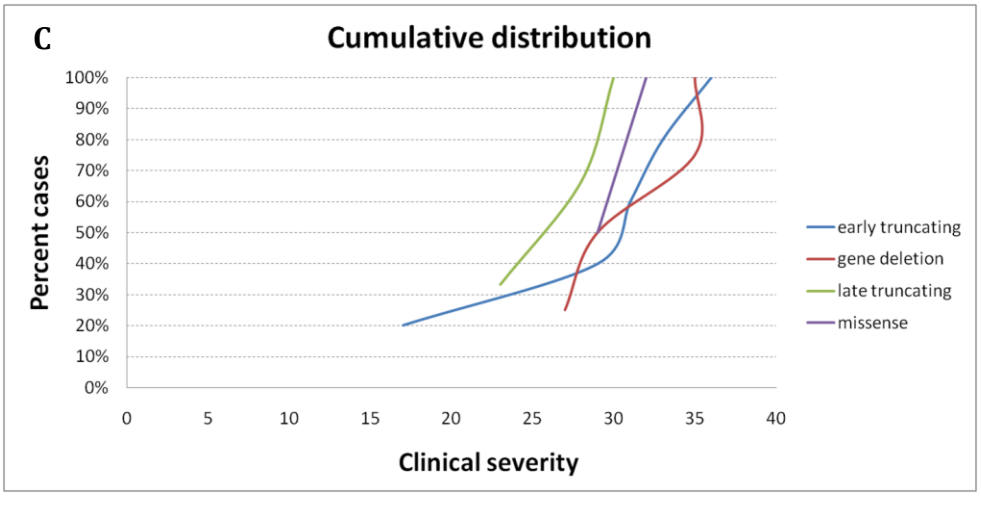


Figure 1 A. Cumulative distribution plots of the patients positive for a *MECP2* mutation. Mutations were grouped based on the four common point mutations (Arg106Trp, Arg133Cys, Thr158Met, Arg306Cys), the four early truncating mutations (Arg168*,

Arg255*, Arg270*, Arg294*), late truncating mutations (LTM), large deletions and all other mutations. Early truncating mutations correspond to mutations interrupting the protein before amino acid 310. Large deletions correspond to deletions including either single exon or the entire gene. **B**. Cumulative distribution of the patients positive for a *CDKL5* mutation. Mutations were grouped based as early truncating mutations (mutations interrupting protein before amino acid 297), late truncating mutations, large deletions (deletions involving either single exons or the entire gene), and missense mutations. **C**. Cumulative distribution of the patients positive for a *FOXG1* mutation. Mutations were grouped based as early truncating mutations (mutations interrupting protein before amino acid 275), late truncating mutations, gene deletions (deletions involving either single exons or the entire gene), and missense mutations.

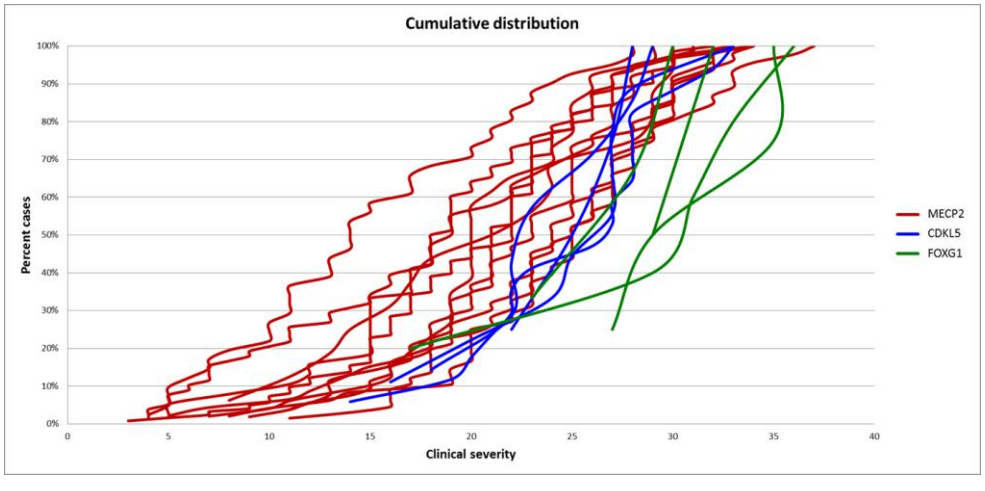


Figure 2 Combined graph illustrating the different clinical severities between *MECP2*, *CDKL5*, and *FOXG1* mutated patients.

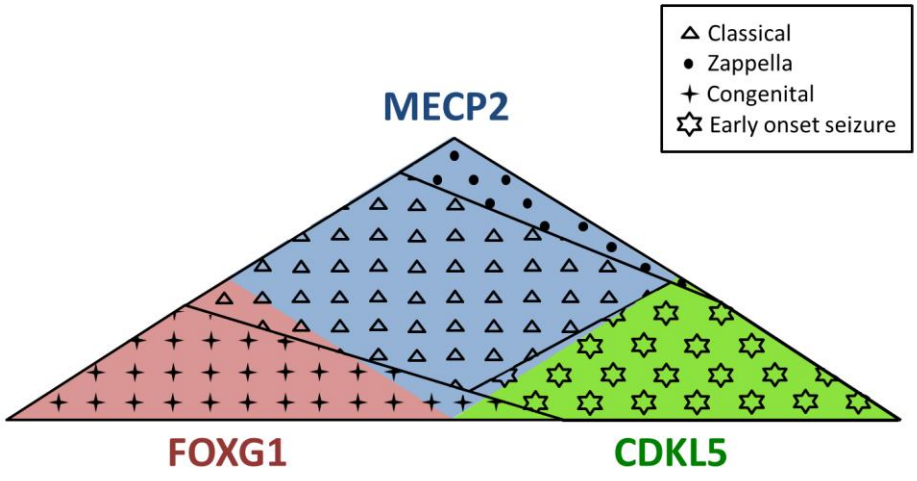


Figure 3 Genotypes and phenotypes in RND. The majority of *MECP2* mutated patients (light blue) have the classical form (triangles), the majority of *CDKL5* mutated patients (green) have the early onset seizure variant (stars) and the majority of *FOXG1* mutated patients (pink) have the congenital form (crosses). Several exceptions to this rule are present: among the *MECP2* mutated patients (light blue) about 6% have the Zappella variant (dots), about 2% has the congenital variant of RTT (crosses), and about 0,4% have the early onset seizure variant (stars); among the *CDKL5* mutated patients (green) about 1,8% patients have the Zappella variant (dots) and another 1,8% have the congenital form (crosses); among the *FOXG1* mutated patients (pink) about 5% of patients has the classical form (triangles) [2].