Genetic analyses of aplastic anemia and idiopathic pulmonary fibrosis patients with short telomeres, possible implication of DNA-repair genes

Elena G. Arias-Salgado,1,2, Eva Galvez,3 Lurdes Planas-Cerezales,4 Laura Pintado-Berninches,1,2, Elena Vallespin,5 Pilar Martinez,5, Jaime Carrillo1, Laura Iarriccio1,2, Anna Ruiz-Llobet,6,7, Albert Catala,6,7, Isabel Badell-Serra,8 Luis I. Gonzalez-Granado,9 Andrea Martin-Nalda,10, Mónica Martínez-Gallo,10 Ana Galera-Miñarro,11 Carmen Rodríguez-Vigil,12 Mariana Bastos-Oreiro,13 Guiomar Perez de Nanclares,14, Virginia Leiro-Fernández,15 Maria-Luz Uriá,10, Cristina Díaz-Heredia,10 Claudia Valenzuela,16, Sara Martín,6, Belén López-Muñiz,17 Pablo Lapunzina,6,18 Julian Sevilla,3,18, María Molina-Molina,4,19 Rosario Perona,1,18† and Leandro Sastre1,18*†

Abstract
Background: Telomeres are nucleoprotein structures present at the terminal region of the chromosomes. Mutations in genes coding for proteins involved in telomere maintenance are causative of a number of disorders known as telomeropathies. The genetic origin of these diseases is heterogeneous and has not been determined for a significant proportion of patients.

Methods: This article describes the genetic characterization of a cohort of patients. Telomere length was determined by Southern blot and quantitative PCR. Nucleotide variants were analyzed either by high-resolution melting analysis and Sanger sequencing of selected exons or by massive sequencing of a panel of genes.

Results: Forty-seven patients with telomere length below the 10% of normal population, affected with three telomeropathies: dyskeratosis congenita (4), aplastic anemia (22) or pulmonary fibrosis (21) were analyzed. Eighteen of these patients presented known pathogenic or novel possibly pathogenic variants in the telomere-related genes TERT, TERC, RTEL1, CTC1 and ACD. In addition, the analyses of a panel of 188 genes related to haematological disorders indicated that a relevant proportion of the patients (up to 35%) presented rare variants in genes related to DNA repair or in genes coding for proteins involved in the resolution of complex DNA structures, that participate in telomere replication. Mutations in some of these genes are causative of several syndromes previously associated to telomere shortening.

Conclusion: Novel variants in telomere, DNA repair and replication genes are described that might indicate the contribution of variants in these genes to the development of telomeropathies. Patients carrying variants in telomere-related genes presented worse evolution after diagnosis than the rest of patients analyzed.

Keywords: Telomere, Dyskeratosis congenita, Pulmonary fibrosis, Aplastic anemia, DNA repair, Telomeropathies

© The Author(s). 2019 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background

Chromosome ends are protected by nucleoprotein structures, the telomeres. In humans, telomere DNA is composed by multiple repetitions of the TTAGGG hexanucleotide [1]. The 3′ end of telomeres is formed by a G-rich single stranded extension that invades the upstream double-stranded region to form telomeric loops (T-, D-loop). Telomere DNA is bound by a protein complex named shelterin [2] and this nucleo-protein structure protects chromosome ends from degradation and is critical for genome stability (recently reviewed in [3]).

Telomeres cannot be completely synthesized during DNA replication, which is known as the end-replication problem and results in progressive shortening of the telomeres as cells proliferate. In most organisms, telomere DNA is elongated after each round of DNA replication by a specialized complex with reverse transcriptase activity known as telomerase [4]. The enzymatic activity of the complex is provided by the TERT (telomerase reverse transcriptase) protein while the RNA component TR (telomerase RNA, encoded by TERC) is used as template. Essential components of the complex are also DKC1 (Dyskeratosis congenita), NOP10, NHP2 and GAR1 proteins, which bind and stabilize TR and are required for telomerase complex assembly [5].

Telomere replication, elongation and protection are impaired by the strong secondary and tertiary structure of their DNA [6]. The high content of Guanines of the telomeres makes them prone to form secondary structures such as G-quadruplexes. In addition, the T- and D-loops impair the access of the replicative machinery. These structures are solved by helicases and structure-specific endonucleases.

In the absence of a functional telomerase complex, telomeres are progressively shortened. One of the main functions of telomere structure is to avoid the recognition of telomeres as damaged DNA which is achieved by inhibitory interactions between shelterin components and DNA-repair proteins [7]. When several telomeres reach critically short length the structure of the nucleo-protein complex can no longer be maintained. In that case, telomeres are recognized as damaged DNA and a response is triggered that can result in cell-cycle arrest, cell apoptosis or senescence [8].

The presence of functional mutations in the genes coding for proteins of the shelterin and telomerase complexes and auxiliary proteins causes a number of diseases named telomeropathies or telomere biology disorders that may affect different organs [9, 10]. Telomeropathies are very heterogeneous diseases depending on the gene mutated and the specific mutations, their penetrance and the existence of anticipation effects. Therefore, patients with the same mutation can present different manifestations. Some patients present severe symptoms at an early age such as those of dyskeratosis congenita (DC) or the related Hoyeraal-Hreidarsson, Resvesz and Coats plus syndromes [11]. Other diseases may appear later such as Aplastic Anemia, (AA) (20–30 years) [12] or pulmonary fibrosis (40–60 years) (PF) [13]. Mutations in 13 different genes have been identified as causative of these diseases but a large number of patients remain genetically undiagnosed [14].

Variants in the gene coding for one helicase, RTEL1 (OMIM 608833) are also present in patients with telomeropathies [15]. Others helicases and nucleases, like SLX4, BLM, WRN and RecQL4, have been shown to participate in telomere preservation [16]. Actually, mutations in some of these genes also result in telomere shortening [16, 17]. The term secondary telomeropathies has been recently proposed for diseases caused by mutations in this group of genes [18].

The reduced number of patients of these rare diseases and the number of putative candidate genes makes their molecular diagnosis challenging. In this article, the molecular analysis of 47 patients treated in Spanish hospitals was performed. Twenty-six of them were diagnosed of DC or AA and twenty-one of PF. The results obtained identified possibly pathogenic variants at the telomere-associated genes TERT, TERC, ACD, CTC1 and RTEL1 in heterozygosis. In addition, a significant proportion of patients presented rare variants in genes coding for proteins involved in DNA repair and in the resolution of complex DNA structures.

Methods

Patients

Patients with severe telomere shortening (percentile < 10) who presented dyskeratosis congenita (DC), aplastic anemia (AA) or pulmonary fibrosis (PF) (sporadic and familial forms) were included. These patients were diagnosed, treated and followed in Spanish reference health care provider centres for these rare diseases. DC was diagnosed in patients with the characteristic mucocutaneous symptoms, nail dystrophy, abnormal skin pigmentation, oral leukoplakia and bone marrow failure. The diagnosis of AA was based on bone marrow and blood cells counts. The diagnosis of PF was established in accordance to international guidelines [19]. The more relevant characteristics of the patients are summarized in Additional file 1: Table S1. All samples were collected after obtaining informed consent for genetic analysis, including for research. Clinical and demographic data were collected.

Genetic analyses

The existence of sequence variants in the TERT, TERC and DKCI genes was determined in six patients by PCR amplification of exons, high resolution melting analyses (HRM) and sequencing of candidate exons as previously described.
Later on, genetic analysis of 41 patients was made by massive parallel sequencing using a panel of genes related to haematological disorders, as shown in Additional file 1: Table S2 (MBFsv1.1 panel). Detailed protocols are presented in Additional file 1: Supplemental Material.

**Telomere length determination**
Telomere length was determined by two different methods, Southern blots of enzymatically digested DNA (Additional file 1: Figure S1) and quantitative PCR. Both methods have been previously used for determination of telomere length from patients with inherited bone marrow failure [21]. Detailed protocols are presented in Additional file 1: Supplemental Material.

**Ethical statement**
All the participants in this study provided written informed consent for genetic analysis. The investigation was conducted in accordance with the ethical standards of the Declaration of Helsinki and the guidelines of the concerned hospitals.

**Results**

**Single nucleotide variants in genes related to telomere biology**
The possible genetic bases of the diseases were studied in patients that presented symptoms related to dyskeratosis congenita (DC), either mucocutaneous symptoms or, more often, aplastic anemia (AA) or sporadic or familial cases of pulmonary fibrosis (PF). Patients with these diseases present significant telomere shortening. According to previous reports [12, 22], patients with telomere length below the 1% of the age-mated population in the case of DC and <10% in AA and PF where included in this study. Beside clinical symptoms, DC/AA and PF patients differed at the age of presentation of the disease, 12.2 ± 12.2 years for DC/AA and 58.3 ± 10.2 years for PF (Additional file 1: Table S1).

First analyses were made in six patients by PCR amplification of exons of the *DKC, TERT* and *TERC* genes using the previously described High-resolution melting and DNA sequencing method [20] (Patients indicated by an asterisk in Tables 1, 4 and Additional file 1: Table S1). Subsequently, taking advantage of newer technologies, DNA samples of 41 patients were analyzed by massive sequencing of a panel of genes related to haematological disorders including all the genes previously associated to telomere-related diseases (Additional file 1: Table S2). Out of these patients, 23 were diagnosed of DC/AA and 18 of PF.

Nucleotide variants were further analysed when their frequency in the general population was lower than one in a thousand. The results obtained for genes previously related to telomere shortening are shown in Table 1. These studies allowed the identification of single nucleotide variants (SNVs) and indels at *TERT, TERC, RTEL1, ACD* and *CTC1* genes in heterozygosis. Histograms of the Sanger sequencing reaction confirming these variants are shown in Additional file 1: Figure S2.

*TERT* variants were found in eight patients (Table 1). AA patient 129 presented two missense variants, p.P65T and p.C931R. The p.P65T variant was inherited from the mother of the patient, which also presented short telomeres. The p.C931R variant was not present in the parents of the patient. AA patient 196 presented in heterozygosis a deletion of two nucleotides at positions c.475_476 that produced a frame shift mutation after the amino acid residue 158. One stop codon will be found 31 residues downstream, resulting in a 189 amino acids-long protein. This variant was inherited from the father of the patient, which presented PF. A son of the patient that also presented this variant, died before birth. AA patient 263 carried in heterozygosis the *TERT* p.V84E variant. His father also carries the variant in heterozygosis and presents PF and short telomeres. A sister that carries the variants is presently unaffected at the age of 33 years (Fig. 1).

PF patient F075 presented two missense variants at *TERT*, p.R447C and p.R742H. The variant p.R742H has been already reported [23] while the p.R447C variant is described for the first time. The mother and a brother of this patient died of PF. PF patient F185 presented the p.S602 L amino acid variant. PF patient F219 presented the p.R671W variant, reported to be pathogenic [24]. PF patient F220 presented an in-frame deletion at amino acids 671_672 that has not been previously described. A variant in a splicing consensus site was found in patient F400.

Variants in *TERC* were found in three patients (Table 1). AA patient 156 presented in heterozygosis a deletion of two bases (r.96_97delCU) that has been previously described [25]. The mother of the patient died with PF and liver cirrhosis, a brother carrying this variant also presented liver cirrhosis and short telomeres. However, a sister that carries the variant is presently healthy at the age of 40 years even if her telomeres are short. PF patient F120 presented the r.269G > A variant, that has not been previously reported. PF patient F253 presented the undescribed r.23G > C variant. PF patient F219 presented the p.P722S variant in *CTC1* that has not been previously described.

Four patients presented rare variants in *RTEL1* in heterozygosis (Table 1). AA patient 90 presented the p.R998* variant recently reported as pathogenic [26]. The father and one sister of the patient carried this variant, and presented short telomeres (Fig. 1). A representative Southern blot of telomeres including relatives of patient 90 is shown in Additional file 1: Figure S1. In addition, patient 90 and mother carried the p.Q929del variant, recently described [26]. AA patient 211 presented the new p.S483I variant and the described p.C1244R variant [27].
patient 211 presented the amino acid variant p.F818S in *PALB2* protein, also involved in telomere maintenance, as will be discussed later. The *RTEL1* variant p.C1244R was inherited from the father and the *RTEL1* variant p.S483I and the *PALB2* variant from the mother and both parents presented average telomere length (Fig. 1). PF patient F001 presented the p.R998* pathogenic variant. This *RTEL1* variant was present in the three analyzed relatives of patient F001, all with short telomeres (Fig. 1). PF patient F020 presented the described p.F988L [15] and the non-described p.R574C variants.

Nucleotide variants in *ACD*, coding for the TPP1 shelterin protein, were observed in two patients. AA patient 145 presented the p.P16Q amino acid variant and AA patient 109 the p.V94I variant, both in heterozygosis. The variant of patient 145 is inherited from the mother that reports several relatives affected of DC. Patients 109 and 145 presented chromosome 7 monosomy.

**Table 1** Mutations found in telomere-related genes

| Gene   | Nucleotide variant | Amino acid variant | Variant reference | Patient | Evolutionary conservation | Functional prediction** | ExAC frequency |
|--------|--------------------|--------------------|-------------------|---------|---------------------------|------------------------|--------------|
| *TERT* | c.193C>A           | p.P65T             | rs5442157565      | 129*    | Non-conserved              | Possibly damaging      | 0.00000      |
|        | c.251T>A           | p.V84E             | Not described     | 263     | Conserved                  | Damaging               | Absent       |
|        | c.475_476delTT     | p.F159Cfs*32       | Not described     | 196*    | Damaging                  | Absent                 |              |
|        | c.2791T>C          | p.C931R            | Not described     | 129*    | Conserved                 | Damaging               | Absent       |
| *TERC* | rs.96_97 delCU     |                    | rs199422267       | 156*    |                           |                        |              |
| *RTEL1*| c.1448G>T         | p.S483I            | Not described     | 211     | Conserved                 | Damaging               | Absent       |
|        | c.3612_3614delCAG  | p.Q929del         | Not described     | 90      | Damaging                  | Absent                 |              |
|        | c.2992C>T          | p.R998*            | rs398123017       | 90      | Conserved                 | Damaging               | 0.00002      |
| *ACD*  | c.4073T>C          | p.C1244R          | rs587777037       | 211     | Conserved                 | Damaging               | 0.00002      |
|        | c.47C>A            | p.F16Q             | Not described     | 145     | Non-conserved              | Possibly damaging      | Absent       |
|        | c.280G>A           | p.V94I             | rs149365469       | 109     | Conserved                 | Pos. dam               | 0.00081      |

*Patients analyzed by exon amplification and Sanger sequencing

**The prediction shown is a summary of the ones made by the programs SIFT ensemble 66, Polyphen-2 v2.2.2, MutationAssessor, release 2, FATHMM v2.3, CADD v1.3 and dbSNV1.1.

2. Pulmonary Fibrosis

| Gene   | Nucleotide variant | Amino acid variant | Variant reference | Patient | Evolutionary conservation | Functional prediction** | ExAC frequency |
|--------|--------------------|--------------------|-------------------|---------|---------------------------|------------------------|--------------|
| *TERT* | c.1339C>T          | p.R447C            | rs1304418053      | F075*   | Non-conserved             | Pos. Dam.              | 0.00000      |
|        | c.1805C>T          | p.S62L             | rs759017540       | F185*   | Conserved                 | Damaging               | 0.00001      |
|        | c.2011C>T          | p.R671W            | rs106050311       | F219    | Not-conserved             | Damaging               | 0.00000      |
|        | c.2013_2015del     | p.R672del          | Not described     | F220    | Damaging                  | Absent                 |              |
|        | c.2225G>A          | p.R742I            | rs727503468       | F075*   | Conserved                 | Damaging               | 0.00000      |
|        | c.2286+1G>A        |                    |                   | F400*   | Damaging                  | Absent                 |              |
| *TERC* | rs.96_97 delCU     |                    |                   |         |                           |                        |              |
| *RTEL1*| c.1720C>T         | p.R574G            | rs3694119645      | F020    | Conserved                 | Damaging               | 0.00004      |
|        | c.2964T>G          | p.F988I            | rs1470145133      | F20     | Conserved                 | Damaging               | 0.00000      |
| *CTCI* | c.2992C>T          | p.R998*            | rs398123017       | F001    | Conserved                 | Damaging               | 0.00002      |
|        | c.2164C>T          | p.P722S            | rs770505786       | F219    | Not-conserved             | Possibly damaging      | 0.00000      |

Single nucleotide variants in genes related to DNA repair and homologous recombination

Massive sequencing identified a number of rare SNVs (frequency lower than one in a thousand) in genes coding for proteins involved in DNA repair by homologous recombination, as shown in Table 2. Among these genes are those coding for the protein kinases ATM (Ataxia-telangiectasia mutated) and ATR (Ataxia telangiectasia and Rad3-related), mutated in the genetic-instability syndromes Ataxia Telangiectasia and Seckel syndrome, respectively. Rare variants in these proteins were found in 35% of the DC/AA and 17% of the PF patients (Table 2). Fanconi Anemia is also due to cytogenetic instability, hypersensitivity to DNA
crosslinking agents and defective DNA repair. This disease can be caused by mutation in genes that code for components of a protein complex involved in DNA repair. Rare variants in several of these genes (FANCA, FANCD2, FANCF, FANCL, FANCM, FANCU) were found in 22% of the DC/AA and PF patients (Table 2). The DNA-associated proteins BRCA1 and BRCA2 (Breast Cancer 1 and 2) also play an important role in DNA break repair and recombination in association with PALB2 (Partner and Localizer of BRCA2). Rare variants in the BRCA1, BRCA2 or PALB2 genes were found in 26% of the DC/AA and 17% of the PF patients (Table 2). RecQL4 is a DNA helicase (RecQ-like, type 4) involved in DNA damage repair and rare variants were found in 17 and 11% of DC/AA and PF patients, respectively. Variants in the gene coding for the structure-specific-endonuclease subunit SLX4, involved in DNA repair and replication, were found in 22% of the DC/AA and 6% of the PF patients (Table 2). Mutations in SLX4 have been associated with Fanconi Anemia (group P). Finally, single rare variants were found in other genes coding for proteins involved in DNA repair such as WRN, BLM, RAD51, RAD51C and NBN in 9 and 17% of the DC/AA and PF patients, respectively. The variants found in these genes for each patient are shown in Table 3. Some of these variants are

Table 2 Frequency of rare variants in genes coding for proteins involved in DNA-Repair

|                | Dyskeratosis/Anemia | Pulmonary Fibrosis |
|----------------|---------------------|--------------------|
| ATM/ATR        | 8/23 (35%)          | 3/18 (17%)         |
| BRCA1, BRCA2, PALB2 | 6/23 (26%)        | 3/18 (17%)         |
| FANCA, −D2, −F, −L, −M, −U | 5/23 (22%)     | 4/18 (22%)         |
| RECQL4         | 4/23 (17%)          | 2/18 (11%)         |
| SLX4           | 5/23 (22%)          | 1/18 (6%)          |
| WRN, BLM, NBN, RAD51, RAD51C | 2/23 (9%)      | 3/18 (17%)         |

Fig. 1 Familiar inheritance of variants in telomere-related genes in patients with aplastic anemia and pulmonary fibrosis. The family of four affected patients is represented in panels (a–d). The probands (indicated by arrows) of families (a–c) are affected of aplastic anemia and that of family (d) of pulmonary fibrosis. The number of the family and the variants analyzed are shown in the upper part of each diagram. In the cases where several variants were studied, they are shown in the heading in the same relative order and colour as in the diagrams. The presence of the variants is represented as colour-filled squares or circles and that of the presence of the reference allele as empty figures. The number assigned to each patient or relative is indicated in the upper, right part of each symbol. The age of each individual, the determined telomere size, in kb, and the corresponding percentage of the age-matched population are shown under each symbol. Relatives affected by pulmonary fibrosis are indicated in the upper left part of their symbol (PF). Patient F001 in panel (d) had a deceased identical twin brother as indicated by joining them with a broken line.
### Table 3 Variants in DNA repair genes

| Gene     | Nucleotide variant | Amino acid variant | Variant reference | Patient   | Evolutionary conservation | Functional prediction* ExAC frequency |
|----------|--------------------|--------------------|-------------------|-----------|---------------------------|----------------------------------------|
| ATM      | NM_0000513         | p.S353T            | rs35963548        | 204       | Conserved                 | 0.00021                                |
|          | NP_0000423         | p.R2355X           | rs2355096         | 69        | Non cons. Pos. damag.     | 0.00086                                |
|          |                    |                    |                   |           |                           |                                        |
| ATR      | NM_0011843         | p.R2297            | rs2229033         | 47        | Conserved                 | 0.00001                                |
|          | NP_0011752         | p.H2391Q           | rs2391914         | 36        | Non cons.                 | 0.00002                                |
|          |                    |                    |                   |           |                           |                                        |
| BRCA1    | NM_0003603         | p.S153S            | rs1800774         | F106      | Damaging                  | 0.00215                                |
|          | NP_0923112         |                    |                   |           |                           |                                        |
| BRCA2    | NM_0000593         | p.N319T            | rs5539572         | F001      | Non cons. Pos. damag.     | 0.00001                                |
|          | NP_0005052         | p.L1019V           | rs5636833         | 90        | Conserved                 | 0.00017                                |
|          |                    |                    |                   |           |                           |                                        |
| PALB2    | NM_0246733         | p.R712V            | rs14158731        | 198       | Non cons.                 | 0.00025                                |
|          | NP_0799512         | p.F811S            | rs37090830        | 211       | Non cons.                 | 0.00004                                |
| FANCA    | NM_0001352         | p.E246K            | rs201992220       | F220      | Conserved                 | 0.00006                                |
| FANCG    | NM_0330843         | p.A1147P           | Not described     | 47        | Damaging                  | Absent                                 |
| FANCL    | NM_0272753         | p.A81S             | rs14650718        | 44        | Non cons.                 | 0.00044                                |
|          | NP_0756211         |                    |                   |           |                           |                                        |
| FANCM    | NM_0246733         | p.K96N             | Not described     | 198       | Conserved                 | Damaging Absent                        |
|          |                   |                    |                   |           |                           |                                        |
| FANCU(XRC2) | NM_054311   | p.R1860H           | rs77535621        | 52        | Conserved                 | Damaging Absent                        |
|          | NM_0543111         |                    |                   |           |                           |                                        |
| FANCU(XRC2) | NM_054311   | p.Y187C            | Not described     | F219      | Conserved                 | Damaging Absent                        |
|          | NM_0543111         |                    |                   |           |                           |                                        |
| SLX4     | NM_0244442         | p.G141W            | rs137976282       | F129      | Conserved                 | 0.00008                                |
|          | NP_1152020         | p.V197A            | rs147827649       | 382       | Conserved                 | 0.00148                                |
|          |                    |                    |                   |           |                           |                                        |
| RECQ4    | NM_0004753         | p.G288A            | Not described     | 181       | Conserved                 | Pos. damag. Absent                     |
|          | NP_0449133         | p.N616S            | rs19965783        | F044      | Conserved                 | Damaging Absent                        |
|          |                    |                    |                   |           |                           |                                        |
| VRN      | NM_0000534         | p.Y1034F           | rs200379049       | 48        | Non cons.                 | 0.00036                                |
|          | NP_0005442         |                    |                   |           |                           |                                        |
| RLM      | NM_0000572         | p.S778C            | rs139610577       | 52        | Conserved                 | Damaging 0.00013                      |
|          | NP_0004831         |                    |                   |           |                           |                                        |
| RAD51C   | NM_0582162         | p.F6L              | rs774685897       | F253      | Non cons.                 | 0.00001                                |
|          | NP_4781231         |                    |                   |           |                           |                                        |

*The prediction shown is a summary of the ones made by the programs SIFT ensemble 66, Polyphen-2 v2.2.2, Mutation Assessor, release 2, FATHMM v2.3, CADD v1.3 and dbscSNV1.1.*
described for the first time as indicated (“Not described” in Table 3).

The variants present in each patient and their clinical evolution from diagnosis are summarized in Table 4. Two patients with previously described DKC1 mutations, 26 and 36 [20], are included. This table includes some variants whose frequency is slightly higher than one in a thousand (indicated in bold letters). One of them is a variant in the gene POT1 (p.Q301H) present at a frequency of 0.002 in the ExAC database. The change is possibly damaging and is present in patient F106. A second variant in NHP2 (p.R101Q) is present at a frequency of 0.002 at the ExAC database. The change is possibly damaging, and has been detected in three patients (F200, F120 and 223). Among the proteins involved in DNA repair, two patients (F20, 109) present the p.T287A variant at the RAD51C gene (frequency of 0.003 in the ExAC database). Other two patients (26 and 134) present the p.A492D variant at the MRE11A gene (frequency 0.002 in ExAC).

The possible correlation between the presence of rare variants in proteins involved in DNA repair or replication, telomere length and clinical manifestation was investigated in the family of several patients (Fig. 2). AA patient 134 carries one variant in the ATR gene not previously described. The variant is inherited from his
mother and grandmother, both relatives have short telomeres. The father and maternal grandfather, that do not carry this variant, present an average telomere size. AA patient 181 present a variant in RECQL4 not described previously. This variant is present in the father with very short telomeres. The mother does not carry

Fig. 2 Familiar inheritance of variants in genes related to DNA repair and replication in patients with aplastic anemia and pulmonary fibrosis. Schematic representation of four families of patients (indicated by arrows) affected by aplastic anemia (a-c) or pulmonary fibrosis (d). The number of each family and the variants studied are shown in the upper part of each panel. The presence of the rare variant is indicated as coloured square or circles and the presence of the reference variant by open figures. In panels (c and d) where several variants were studied, their relative position in the diagrams and colour are indicated in the upper part. The number of each family member and its possible pulmonary fibrosis diagnosis (PF) is indicated in the upper part of each symbol. The age, telomere length (kb) and the corresponding percentage of the age-matched population are indicated in the lower part of each symbol.
this variant and has longer telomeres even that below the average of the population (10%).

AA patient 198 presents a novel variant in ATM and a rare variant in RAD50. The ATM variant is inherited from the mother and the RAD50 p.R850Q variant from the father (Fig. 2). The father present telomeres of average size and the mother of size bellow the average. However, the son presents relatively shorter telomeres than the parents.

The PF patient F106 carries a rare variant in BRCA1 and a relatively rare variant in POT1 (frequency 0.002). The patient has four brothers and three have been also diagnosed of PF. Two of them carry the same variants in POT1 and BRCA1 and short telomeres. A sister carries only the BRCA1 variant and very short telomeres although is presently asymptomatic. Two of the brothers carry exclusively the POT1 mutation and present longer telomeres (25–50% of the population) although one of them has developed pulmonary fibrosis. The patient has a daughter, which inherited the BRCA1 variant and present very short telomeres. Two nephews and a niece carry the BRCA1 variant and have short telomeres. One niece inherited the POT1 change and presents short telomeres.

**Discussion**

The genetic variants present in patients with symptoms of DC, AA or PF have been studied in a population of patients from Spanish hospitals. Heterozygous variants were found in genes previously related to telomeropathies in 17 patients, 8 diagnosed of DC/AA and 9 of PF (Table 1). Two variants in ACD were found in the only two AA patients that also presented chromosome 7 monosomy, previously associated to the presence of short telomeres in AA patients as a consequence of chromosomal instability [28].

Several of the variants found had been reported previously as pathogenic. The TERT variants p.R671W and p.R742H have been described associated to PF [23, 24]. The TERC variant r.96_97delCU was described in a DC patient transmitted with a dominant pattern of inheritance [25]. The RTEL1 variants p.Q929del [26] p.F988L [29], p.R998* [26] and p.C1244R [27] were described in DC patients. The RTEL1 p.R574C variant has been described as a clinical variant. The other variants found in this study have not been previously reported and according to the actual guidelines should be considered as variants of unknown significance (VUS). However, we speculate that some of them could be pathogenic because they are closely related to described pathogenic variants, as schematically shown in Fig. 3. This is the case of TERT p.P65T variants since the p.P65A variant has been found in patients of acute myeloblastic leukemia (AML) [26]. The TERT p.V84E variant is close to the p.R83P variant, found in patients with AA and myelodisplastic syndrome [30]. In addition, the V84 residue is conserved through evolution and the V > E change is considered as damaging. Other variants produced changes potentially damaging in conserved residues located in functional domains of the proteins like the TERT variants p.F159fsX32 and p.S602L that affect residues located in the conserved N-terminal domain (Fig. 3). The p.F159fsX32 variant would produce a truncated protein and a situation of functional hemizygosis. The variants p.R672del and p.C931R are located in the reverse transcriptase and C-terminal domains of TERT, respectively (Fig. 3). In the case of the RTEL1 gene, the p.S483I variant in a conserved residue is close to the p.P484L variant associated to PF [31].

The gene panel used for the analyses of the DNA of 41 patients by massive sequencing detected the presence of numerous variants in genes coding for proteins.

**Fig. 3** Schematic representation of the location of TERT variants on the conserved and functional domains of the protein. Conserved regions of TERT are represented as filled boxes and coloured according to their participation in the protein functional domains (shown in the lower part of the scheme). Regions of the N-terminal extension, involved in RNA binding, are coloured in blue, Regions of the reverse transcriptase domain (RT-domain) are shown in cyan and those of the C-terminal extension in green. The location of the new amino acid variants described in the article is shown in the upper part of the figure. Two previously described variants close to new variants (P65A and R83P) are shown in the lower part of the figure. The location of two variants found in this project and previously described as pathogenic are shown in the lower part inside boxes.
involved in DNA repair such as ATM, ATR, BRCA1, BRCA2 and several genes of the FANC complex, mutated in Fanconi Anemia. Variants in genes coding for DNA helicases and nucleases involved in the resolution of DNA complex structures, such as SLX4, RECQL4, WRN, BLM, RAD51, RAD51C or NBN were also found. Variants in these genes have not been related to telomeropathies previously, with the exception of SLX4 [32].

The proteins ATM and ATR play a signalling role in DNA double-strand break repair pathways and their continued activation can result in cell cycle arrest, replicative senescence or apoptotic cell death. However, specific interactions with shelterin protein, ATM/TRF2, and ATR/TRF1, inhibit their activation to preserve telomere structure. Therefore, it is conceivable that functional variants at ATM and ATR could affect telomere structure and stability. Indeed, cells of Ataxia telangiectasia patients with homozygous mutations in ATM present telomere shortening and increase frequency of chromosome end-to-end fusions (CEFs) [33].

BRCA1 and BRCA2 proteins also participate in DNA damage repair by homologous recombination and non-homologous end joining [34] and play a protective role at telomeres [35]. BRCA1/2 mutations result in shortening of the telomere single-stranded overhang and increase telomere length variability [36]. BRCA2 heterozygous cell lines present frequently CEFs [37]. Germline mutations in BRCA1 and BRCA2 are associated with greatly increased frequency of breast and ovarian cancers and telomere shortening was associated with genetic anticipation in hereditary breast cancer [38].

Proteins mutated in Fanconi Anemia (FA) patients are involved in repair of DNA-inter-strand crosslinks and have been also involved in telomere maintenance (reviewed by Sarkar and Liu [39]). FA patients are reported to present increased CEFs and overall shorter telomeres [40, 41].

Replication of telomeric DNA presents specific challenges due to structure of the DNA with unconventional regions including T loop and G quadruplexes and requires the participation of a large number of proteins (reviewed by Martinez and Blasco [6]). Among them are proteins with DNA helicase activity, like RTEL1, BLM, WRN or RecQL4, endonuclease complexes, like the one formed by SLX4, SLX1, MUS81 and XPF and recombinases like Rad51 and Rad51C.

Mutations in these genes cause diseases with some symptoms resembling those of telomeropathies, as mentioned in the Background section. Many of these diseases show telomere fragility [42], telomere loss and/or increased CEFs [43]. These proteins are located to telomeres through interactions with TRF1 of TRF2 shelterin proteins [44–48]. In addition, TERT expression in cells isolated from patients of these diseases rescue some of their alterations [49, 50]. Variants in SLX4 have been found with increased frequency in AA patients [32].

The existence of overlapping symptoms in diseases caused by mutations in genes coding for proteins involved in telomere-maintenance and in DNA-repair and replication proteins have been recently reviewed [18]. In this article we describe the presence of rare variants in many of these genes in patients that present short telomeres and telomeropathies such as DC, AA and PF. We consider the possibility that the presence of one or more variants in these genes in heterozygosis, some times in addition to variants in telomere maintenance genes, could have a cumulative pathogenic effect. This association could be extensive to some more frequent variants in genes involved in primary telomeropathies, as those observed in NHP2 or POT1. This possible pathogenic association need to be confirmed by functional studies and analyzing more extensive cohorts of patients.

The small number of patients analyzed prevents a statistical analysis of the possible genotype-phenotype correlation. There seems to be, however poor outcome in patients that carry variants in genes related to telomere biology, as shown in Table 4. For example, nine AA patients carried variants in telomere genes and five of them died in the course of the study and the other four required hematopoietic stem cell transplantation (HSCT). In contrast, only one patient died and three required HSCT of the sixteen that presented short telomeres but not rare variants in telomere related genes. Similar evolution was observed in PF patients. Nine carried mutations in telomere genes and four died and four required lung transplant. Of the 12 patients without mutations in these genes, two died and three required lung transplant. In particular, the two patients that carried two TERT variants evolved very quickly and both died in a few months after the diagnosis of the disease. One of these patients was diagnosed of AA and the other of PF. Among the patients with worst evolution without mutations in telomerase genes, three presented variants in FANCM (two died and one required HSCT). However, the analysis of a larger cohort of patients will be required to confirm or discard these proposals.

Conclusions

Using massive sequencing of a panel of genes and Sanger sequencing of selected exons we identified novel variants in telomere biology genes in a series of patients of DC, AA and PF form Spanish hospitals. In additions we describe the existence of frequent rare variants in genes involved in DNA damage response and the resolution of complex DNA structures in this population of patients. Achieving a molecular diagnosis from patients of these telomere biology diseases would be important for accurate genetic counselling and effective treatment of the patients.
Additional file

Additional file 1: Supplemental methods. Table S1A. Clinical characteristics of the Dyskeratosis congenita, Aplastic anemia patients. Table S1B. Clinical characteristics of Pulmonary fibrosis patients. Table S2. Genes included in the panel used for massive sequencing. Table S3. Description of the control population used for the estimation of normal mean, telomere length and percentile for each age-range. Figure S1. Representative Southern blot used for telomere length determination. Figure S2. Histograms of the sequences obtained by Sanger sequencing of the regions of genes related to telomere biology that presented SNVs or indels. (DOC 3474 kb)

Acknowledgments
We are grateful to AFFPFI (Asociación de Familiares y Enfermos de Fibrosis Pulmonar Idiopática) for patients’ support. We acknowledge support of the publication fee by the CSIC Open Access Publication Support initiative through its Unit of Information Resources for Research (URICI).

Funding
Funded by grants PI14-01495 and PI17–01401 (Fondo de Investigaciones Sanitarias, Instituto de Salud Carlos III, Spain) supported by FEDER funds and by one ACCI project from CIBERER and one grant to the FPI cohort from CIBERES.

Availability of data and materials
Additional data and materials can be provided upon request to the authors.

Authors’ contributions
Conception and design of the work: EGA-S, JS, MM-M, RP, LS. Generation of DNA sequencing data: EGA-S, EG, LP-B, EV, PM, JC, LI, PL, RP, LS. Clinical diagnosis and follow up of the patients: LP-C, A-RL, AC, IB-S, LIG-G, AM-N, MW-G, AMG-M, CRV, MB-O, GPN, VL-F, M-LL, CD-H, CV, SM, BL-M. Data analysis and interpretation: EGA-S, EG, JC, JS, MM-M, RP, LS. Critical revision of the article: All the authors. All authors approved the final version of the manuscript to be published.

Ethics approval and consent to participate
This study was conducted in accordance with the principles of the Helsinki Declaration and the guidelines of the concerned hospitals.

Consent for publication
All the participants in this study provided written informed consent for genetic analysis, including for research.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Instituto de Investigaciones Biomedicas CSIC/UAM, IDIPaz, Arturo Dupereier, 4, 28029 Madrid, Spain. 2Advanced Medical Projects, Madrid, Spain. 3Hospital Niño Jesús, Hematologia y Hemoterapia, Madrid, Spain. 4ILD Unit Pneumology Department, University Hospital of Bellvitge, IDIBELL, University of Barcelona, Barcelona, Spain. 5Institute of Medical and Molecular Genetics (INGEMMW), Hospital Universitario La Paz, Madrid, Spain. 6Pediatric Hematology and Oncology Department, Hospital Sant Joan de Déu, University of Barcelona, Barcelona, Spain. 7Institut de Recerca Pediàtrica Hospital Sant Joan de Déu (IRP-HSJD), Esplugues de Llobregat, Barcelona, Spain. 8Hospital de la Santa Creu i Sant Pau, Barcelona, Spain. 9Hospital 12 de Octubre, Madrid, Spain. 10Hospital Universitario Virgen de la Arrixaca, Murcia, Spain. 11Hospital Miguel Servet, Zaragoza, Spain. 12Hospital Universitario Gregorio Marañón, IlSGM, Madrid, Spain. 13Molecular (Epi)Genetics Laboratory, BioAraba National Health Institute, OSI Araba University Hospital, Vitoria-Gasteiz, Spain. 14Pneumology Department, Hospital Alvaro Cunqueiro, Complexo Hospitalario Universitario de Vigo, Neumovigil+ Research Group, Vigo Biomedical Research Institute (IBV), Barcelona, Spain. 15Hospital de La Princesa, Madrid, Spain. 16Hospital Infanta Leonor, Madrid, Spain. 17CIBER of enfermedades raras (CIBERER), Madrid, Spain. 18CIBER of Respiratory diseases (CIBERES), Barcelona, Spain.

Received: 31 October 2018 Accepted: 12 March 2019

References
1. McEIllogg R, Wellinger RJ. The terminal DNA structure of mammalian chromosomes. EMBO J. 1997;16(12):3705–14.
2. de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. 2005;19(18):2100–10.
3. Maciejowski J, de Lange T. Telomeres in cancer: tumour suppression and genome instability. Nat Rev Mol Cell Biol. 2017;18(3):175–86.
4. Schmidt JC, Cech TR. Human telomerase: biogenesis, trafficking, recruitment, and activation. Genes Dev. 2015;29(11):1095–105.
5. Egan ED, Collins K. Specificity and stoichiometry of subunit interactions in the human telomerase holoenzyme assembled in vivo. Mol Cell Biol. 2010; 30(11):2775–86.
6. Martinez P, Blasco MA. Replicating through telomeres: a means to an end. Trends Biochem Sci. 2015;40(9):504–15.
7. KarJeder J, Hoke K, Mirzaeva OK, Bakkenist C, Kastan MB, Petriti JH, et al. The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. PLoS Biol. 2004;2(8):e240.
8. d’Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, von Zglinicki T, et al. A DNA damage checkpoint response in telomere-initated senescence. Nature. 2003;426(6963):194–8.
9. Savage SA. Human telomeres and telomere biology disorders. Prog Mol Biol Transl Sci. 2014;125:41–66.
10. Bertiuc AA. The molecular genetics of the telomere biology disorders. RNA Biol. 2016;13(8):9596.
11. Perona R, Iarriccio L, Pintado-Berninches L, Rodriguez-Centeno J, Manguan-Garcia C, Garcia E, et al. Molecular diagnosis and precision therapeutic approaches for telomere biology disorders. In: Larramendi M, Solonenski S, editors. Telomeres: INTECH. 2016. p. 77–117.
12. Townley DM, DumiItiu B, Young NS. Bone marrow failure and the telomeropathies. Blood. 2014;124(18):2775–83.
13. Armanios M. Telomerase and idiopathic pulmonary fibrosis. Mutat Res. 2012; 730(1–2):52–8.
14. Barbaro P, Ziegler DS, Reddel RR. The wide-ranging clinical implications of the short telomere syndromes. Intern Med J. 2016;46(4):393.
15. Walne AJ, Vullaylam T, Kirwan M, Plagnol V, Dokal I. Constitutional mutations in RTEL1 cause severe dyskeratosis congenita. Am J Hum Genet. 2013;92(3):448–53.
16. Doksani Y, de Lange T. The role of double-strand break repair pathways at functional and dysfunctional telomeres. Cold Spring Harb Perspect Biol. 2014;6(12):a016576.
17. Alter BP, Giri N, Savage SA, Rosenberg PS. Telomere length in inherited bone marrow failure syndromes. Haematologica. 2014;100(1):49–54.
18. OpreSko PL, Shay JW. Telomere-associated aging disorders. Aging Res Rev. 2016;33:52–66.
19. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. Am J Respir Crit Care Med. 2011; 183(6):788–824.
20. Canillo J, Martinez P, Solera J, Moratalla C, González A, Manguan-Garcia C, et al. High-resolution melting analysis for the identification of novel mutations in DKC1 and TERT genes in patients with dyskeratosis congenita. Blood Cells Mol Dis. 2012;49(3–4):140–6.
21. Gadalla SM, Cavthon R, Giri N, Alter BP, Savage SA. Telomere length in blood, buccal cells, and fibroblasts from patients with inherited bone marrow failure syndromes. Aging (Albany NY). 2012;3:523-66.
22. Alter BP, BaeIocher GM, Savage SA, Chancok SJ, Weiskler BB, WIlner JP, et al. Very short telomere length by flow fluorescence in situ hybridization identifies patients with dyskeratosis congenita. Blood. 2007;110(5):1439–47.
23. Petrovski S, Todd JL, Durheim MT, Wang Q, Chien JW, Kelly FL, et al. An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. Am J Respir Crit Care Med. 2017;196(1):82–93.
