

**Germline intergenic duplications at Xq26.1 underlie  
Bazex-Dupré-Christol syndrome, an inherited basal cell carcinoma  
susceptibility condition**

**Authors:** Yanshan Liu<sup>1,†</sup>, Siddharth Banka<sup>2,3,†,\*</sup>, Yingzhi Huang<sup>1,†</sup>, Jonathan Hardman-Smart<sup>4,5</sup>, Derek Pye<sup>4</sup>, Antonio Torrelo<sup>6</sup>, Glenda M. Beaman<sup>2,3</sup>, Marcelo G. Kazanietz<sup>7</sup>, Martin J Baker<sup>7</sup>, Carlo Ferrazzano<sup>8</sup>, Chenfu Shi<sup>8</sup>, Gisela Orozco<sup>8</sup>, Stephen Eyre<sup>8</sup>, Michel van Geel<sup>9,10</sup>, Anette Bygum<sup>11,12</sup>, Judith Fischer<sup>13</sup>, Zosia Miedzybrodzka<sup>14,15</sup>, Faris Abuzahra<sup>16</sup>, Albert Rübben<sup>17</sup>, Sara Cuvertino<sup>2</sup>, Jamie M. Ellingford<sup>2,3</sup>, Miriam J. Smith<sup>2,3</sup>, D. Gareth Evans<sup>2,3</sup>, Lizelotte J.M.T Weppner-Parren<sup>18</sup>, Maurice A.M. van Steensel<sup>19,20</sup>, Iskander H. Chaudhary<sup>21</sup>, D. Chas Mangham<sup>22</sup>, John T. Lear<sup>4,23</sup>, Ralf Paus<sup>4,24,25</sup>, Jorge Frank<sup>26,§</sup>, William G. Newman<sup>2,3,§,\*</sup>, Xue Zhang<sup>1,§,\*</sup>

**Affiliations:**

<sup>1</sup>McKusick-Zhang Center for Genetic Medicine, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100005, China.

<sup>2</sup>Division of Evolution, Infection and Genomics, Faculty of Biology, Medicine and Human Sciences, University of Manchester, Manchester, M13 9PL, United Kingdom.

<sup>3</sup>Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust, Manchester, M13 9WL, United Kingdom.

<sup>4</sup>The Centre for Dermatology Research, University of Manchester, MAHSC, and National Institutes of Health Biomedical Research Center, Manchester, M13 9PL, United Kingdom.

<sup>5</sup>St John's Institute of Dermatology, Kings College London, London, WC2R 2LS, United Kingdom.

<sup>6</sup>Department of Dermatology, Hospital Infantil Universitario Niño Jesús, 28009 Madrid, Spain.

<sup>7</sup>Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States.

<sup>8</sup>Centre for Genetics and Genomics Versus Arthritis Division of Musculoskeletal and Dermatological Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, M13 9PL, United Kingdom.

<sup>9</sup>Department of Dermatology, University Hospital Maastricht, 6229 Maastricht, The Netherlands.

<sup>10</sup>GROW School for Oncology and Developmental Biology, Maastricht University Medical Center +, 6229 Maastricht, The Netherlands.

<sup>11</sup>Department of Clinical Genetics, Odense University Hospital, 5230 Odense, Denmark

<sup>12</sup>Hospital Clinical Institute, University of Southern Denmark, 5230 Odense, Denmark.

<sup>13</sup>Institute of Human Genetics, Medical Center, University of Freiburg,  
79106 Freiburg, Germany.

<sup>14</sup>School of Medicine, Medical Sciences, Nutrition and Dentistry, University  
of Aberdeen, Aberdeen AB25 2ZD, United Kingdom.

<sup>15</sup>Medical Genetics Department, NHS Grampian, Foresterhill, Aberdeen,  
AB25 2ZD, United Kingdom.

<sup>16</sup>Department of Dermatology, Zaandam Medical Center, 1502 Zaandam,  
The Netherlands.

<sup>17</sup>Department of Dermatology and Allergology, University Hospital of  
RWTH Aachen, 52062 Aachen, Germany.

<sup>18</sup>Department of Dermatology, Jeroen Bosch Hospital, 5223  
's-Hertogenbosch, the Netherlands.

<sup>19</sup>Skin Research Institute of Singapore, Agency for Science, Technology  
and Research (A\*STAR), Singapore 138543, Singapore.

<sup>20</sup>Lee Kong Chian School of Medicine, Nanyang Technological University  
(NTU), Singapore 636921, Singapore.

<sup>21</sup>Department of Pathology, Royal Liverpool University Hospital, Liverpool,  
L7 8XP, UK

<sup>22</sup>Adult Histopathology, Laboratory Medicine, Manchester University NHS  
Foundation Trust, Health Innovation Manchester, Manchester, M13 9WL,  
United Kingdom.

<sup>23</sup>Department of Dermatology, Salford Royal NHS Foundation Trust,  
Manchester, M6 8AD, United Kingdom.

<sup>24</sup>Dr Phillip Frost Department of Dermatology & Cutaneous Surgery,  
University of Miami Miller School of Medicine, Miami, FL 33125, United  
States.

<sup>25</sup>Monasterium Laboratory, Nano-Bioanalytik Zentrum, D - 48149 Münster,  
Germany

<sup>26</sup>Dept. of Dermatology, Venereology and Allergology, University Medical  
Center Göttingen, 37075 Göttingen, Germany

†, § Equal contribution

**\*Correspondence to**

**Xue Zhang**, McKusick-Zhang Center for Genetic Medicine, Institute of  
Basic Medical Sciences, Chinese Academy of Medical Sciences and  
Peking Union Medical College, Beijing 100005, China. Phone:  
(+86)-10-65105154. E-mail: [xuezhang@pumc.edu.cn](mailto:xuezhang@pumc.edu.cn);

**William G. Newman**, Division of Evolution, Infection and Genomics,  
Faculty of Biology, Medicine and Human Sciences, University of  
Manchester, Manchester, M13 9PL, United Kingdom. E-mail:  
[william.newman@manchester.ac.uk](mailto:william.newman@manchester.ac.uk)

**Siddharth Banka**, Division of Evolution, Infection and Genomics, Faculty  
of Biology, Medicine and Human Sciences, University of Manchester,

Manchester, M13 9PL, United Kingdom. E-mail:

Siddharth.Banka@manchester.ac.uk

## Abstract

**Background:** Bazex-Dupré-Christol syndrome (BDCS; MIM301845) is a rare X-linked dominant genodermatosis characterized by follicular atrophoderma, congenital hypotrichosis and multiple basal cell carcinomas (BCCs). Previous studies have linked BDCS to an 11.4 Mb interval on chromosome Xq25-27.1. However, the genetic mechanism of BDCS remains an open question.

**Methods:** To investigate the genetic etiology of BDCS, we ascertained eight families with individuals affected with BDCS (F1-F8). Whole exome (F1 and F2) and genome sequencing (F3) were performed to identify putative disease-causing variants within the linkage region. Array-comparative genomic hybridization and quantitative PCR were used to explore copy number variations (CNV) in BDCS families, followed by long-range gap-PCR and Sanger sequencing to amplify duplication junction and define the precise head-tail junctions, respectively. Immunofluorescence was performed in hair follicles, BCCs and trichoepitheliomas from BDCS patients and sporadic BCCs to detect the expression of corresponding genes. The *ACTRT1* variant (p.Met183Asnfs\*17), previously proposed to cause BDCS, was evaluated with allele frequency calculator.

**Results:** In eight BDCS families, we identified overlapping 18-135kb duplications (six inherited and two *de novo*) at Xq26.1, flanked by *ARHGAP36* and *IGSF1*. We detected *ARHGAP36* expression near the control hair follicular stem cells compartment, and found increased *ARHGAP36* levels in hair follicles in telogen, BCCs and trichoepitheliomas from patients with BDCS. *ARHGAP36* was also detected in sporadic BCCs from individuals without BDCS. Our modelling showed the predicted *ACTRT1* variants maximum tolerated

minor allele frequency in control populations to be orders of magnitude higher than expected for a high-penetrant ultra-rare disorder, suggesting loss-of-function of *ACTRT1* is unlikely to cause BDCS.

**Conclusions:** Our data support the pathogenicity of intergenic duplications at Xq26.1, most likely leading to dysregulation of *ARHGAP36*, establish BDCS as a genomic disorder, and provide a potential therapeutic target for both inherited and sporadic BCCs.

**Keywords:** Bazex-Dupré-Christol syndrome, basal cell carcinomas, copy number variation, *ARHGAP36*

## 1 **Background**

2 Bazex-Dupré-Christol syndrome (BDCS, also called Bazex syndrome or follicular  
3 atrophoderma and basal cell carcinomas; MIM 301845) is a rare X-linked disorder  
4 characterized by congenital hypotrichosis, follicular atrophoderma (seen as “ice pick”  
5 marks, usually on the dorsum of hands and feet) and susceptibility to develop basal cell  
6 nevi and basal cell carcinomas (BCCs). The nevi and BCCs generally occur on  
7 sun-exposed areas, including the head, neck and face from the second decade of life [1].  
8 Other features reported in some individuals with BDCS include persistent milia,  
9 hyperpigmented macules, hypohidrosis, and trichoepitheliomas (benign tumors arising  
10 from basal cells in the hair follicles, which rarely transform to BCCs) [2-9].

11 Clinically, BDCS overlaps with basal cell nevus syndrome (Gorlin syndrome; MIM  
12 109400), which predisposes to multiple BCCs and is caused by heterozygous germline  
13 variants in *PTCH1* [10, 11] or *SUFU* [12]. Both genes encode members of the hedgehog  
14 signaling pathway, and variants in them result in dysregulated overexpression of Gli  
15 transcription factors. BCC is the most common skin cancer [13], with somatic mutations  
16 in genes encoding key components of Hedgehog-Patched-Gli signalling often present in  
17 sporadic BCCs [14]. BCCs arise from hair follicle stem cells and/or other epithelial stem  
18 cells reprogrammed to a follicular differentiation [15, 16]. Of note, X-linked inheritance  
19 pattern is unusual for a cancer predisposition syndrome. The ‘two-hit model’ for familial  
20 cancers caused by loss of heterozygosity of the remaining functional allele in an  
21 individual with a germline loss-of-function variant is unlikely to occur for an X-linked  
22 tumor suppressor gene. Hence, we hypothesized that BDCS is caused by genetic



1 variants resulting in altered Hedgehog-Patched signaling or -Gli activity in follicular stem  
2 cells via a mechanism that does not result in direct loss of function.

### 3 **Methods**

#### 4 **Patient ascertainment**

5 The procedures followed were in accordance with the ethical standards of the  
6 responsible committee on human experimentation (institutional and national), and written  
7 informed consents to take part in the present study, as well as to have the results of this  
8 work published were obtained from the participants. We identified patients and families  
9 diagnosed with BDCS based on clinical history and examination.

#### 10 **Sanger sequencing**

11 A single primer pair (ACTRT1-F: TAGGTATGATTTGCTTTCCTTGGC, ACTRT1-R-R:  
12 CAACCTAAAGATTCATGACATGACTC) was designed to amplify the full length of  
13 ACTRT1, which encompassed the single coding exon and its 5' and 3' untranslated  
14 regions (UTR). At least one affected individual from each family underwent Sanger  
15 sequencing.

16 All coding exons of *ARHGAP36* (exons 2-12) were amplified in the panel of lymphocyte  
17 DNA samples from individuals with clinical diagnoses of Gorlin syndrome without causal  
18 variants in *PTCH1* or *SUFU* and underwent Sanger sequencing on an ABI3730xl DNA  
19 Analyser (Life Technologies, Paisley, UK).

#### 20 **Exome and Genome sequencing**

21 Whole exome sequencing was performed by using SureSelect kits (Human All Exon  
22 50Mb for Families 1 and 2 and Human All Exon v.5 for Family 3) (both Agilent, Santa

1 Clara, CA, USA) were used for whole exome sequencing. Paired-end sequencing  
2 (~100bp) was performed on Illumina HiSeq2000 (Family 1 and 2) or HiSeq2500  
3 platforms (Family 3). For the first two families, a minimum of 4.3 Gb of high-quality  
4 mappable data was generated, yielding a mean depth of coverage of 40-fold and 84% of  
5 target bases sequenced at 10x coverage. A minimum of 4.5 Gb of sequence was  
6 generated, yielding a mean depth of coverage of 80-fold and 95% of target bases  
7 sequenced at 20x coverage. The sequence data were mapped to the human reference  
8 genome (hg19) using Burrows Wheeler Aligner [17]. Variant calling was performed using  
9 the GenomeAnalysisToolKit-v2.4.7. (GATK) software [18]. Genome sequence data was  
10 generated by Complete Genomics (Mountain View, California, USA) as described  
11 previously [19]. Bioinformatics (alignment to the hg19 reference genome, local de novo  
12 assembly and variant calling) was performed using version 2.5 of the Complete  
13 Genomics pipeline [20].

#### 14 **Array-comparative genomic hybridisation (a-CGH)**

15 A-CGH was performed using Affymetrix SNP6.0 microarray as described previously [21].

#### 16 **Real-time quantitative PCR (qPCR)**

17 To validate the genomic duplications detected by aCGH and determine the boundaries  
18 of duplications in the three families and other families, qPCR was performed as  
19 previously described [22]. For families 1, 2 and 7, the primers for qPCR were designed  
20 according to aCGH results. For the other families, primer pair XQM located in the middle  
21 of the shared duplicated region (chrX:130348186-130348315) was designed. The  
22 quantification of the target regions was normalized to an assay from chromosome 21.  
23 The relative copy number (RCN) was determined with the comparative  $\Delta\Delta\text{CT}$  method,  
24 using DNA from a normal male as the calibrator. All assays were repeated three times. A

1 ~2-fold RCN indicated duplication in samples from males and ~3-fold RCN for samples  
2 from females. Primers used for each family were listed in Supp Table 3.

### 3 **Haplotype analysis**

4 Four common single nucleotide polymorphisms (SNPs) (rs62603806, rs4240127,  
5 rs5932866, rs12559533) within the duplicated region were amplified with primer pair  
6 (SNP-F: GCACAGATGATTATGTCTGTTCC, SNP-R:  
7 CTGTCCCTACTTAGTAAATCGAG) and Sanger sequenced to generate haplotypes in  
8 the male patients of F1, F3, and F8.

### 9 **Long-range Gap-PCR**

10 A series of qPCR primers was designed to walk through to refine the duplication  
11 boundaries. Once the boundaries were estimated by qPCR, the forward primer from the  
12 centromeric side of the duplication region that was nearest to the determined boundaries,  
13 as well as the reverse primer from the telomeric side were chosen to perform long-range  
14 gap-PCR in order to amplify the duplication junctions. F3, F5 and F8 used the same  
15 primers pair as F1. Sanger sequencing was conducted to define the precise breakpoints.  
16 Primers used for each family were listed in Supp Table 3.

### 17 **Population screening**

18 Primer pair XQM, as described above, was used to determine the frequency of  
19 duplications at the disease associated locus by qPCR. The criteria defining duplication  
20 was the same as above.

### 21 **Immunohistochemistry**

22 Immunohistochemistry was performed on 7um frozen sections by fixation in chilled  
23 paraformaldehyde (4% v/v) for 10 minutes. After washing with phosphate buffered saline

1 (PBS), sections were permeabilised with PBS containing 0.1% Triton X100 for ten  
2 minutes. Primary antibodies were incubated over night at 4°C at a concentration of 1:200  
3 (ARHGAP36, HPA002064, Sigma, Dorset, UK), 1:20 (IGSH, HPA035582, Sigma) or  
4 1:100 (ACTRT1, HPA003119, Sigma). For dual stains cytokeratin 15 (Abcam, ab80522,  
5 Cambridge, UK), at a concentration of 1:500 was added and incubated overnight.  
6 Antibodies were detected by incubating with an Alexafluor goat anti mouse/rabbit  
7 secondary (Sigma) antibody at a concentration of 1:200 for 45 minutes. Sections were  
8 counterstained with DAPI.  
9 For tumor tissue, paraffin sections were de-waxed using xylene and rehydrated using  
10 decreasing concentrations of ethanol (100%-50%). Antigen retrieval was performed by  
11 boiling in 10mM citrate buffer (pH 6) for 20 minutes. Following this the protocol as above  
12 was followed with P63 (Abcam-ab735) at a concentration of 1:100.

### 13 **Allele frequency modelling**

14 Allele frequency modelling of *ACTRT1* was performed using the allele frequency  
15 calculator from <http://cardiodb.org/allelefrequencyapp> [23]. The maximum tolerated  
16 reference allele count (0.95 CI) for the *ACTRT1* NM\_138289.3:c.547dup  
17 (p.(Met183Asnfs\*17)) variant was calculated in 204,684 in alleles  
18 ([https://gnomad.broadinstitute.org/variant/X-127185638-A-AT?dataset=gnomad\\_r2\\_1](https://gnomad.broadinstitute.org/variant/X-127185638-A-AT?dataset=gnomad_r2_1)).  
19 The adjustable parameters were – estimated population prevalence of BDCS (1 in106)  
20 (Orpha.net); allelic heterogeneity of 33% (BDCS in 2 out of 6 families in Bal et al was  
21 attributed to this variant); and penetrance of 100% (based on description of families in  
22 the literature. Next, to account for potential inaccuracies in previous estimates, we  
23 modelled the maximum tolerated reference allele counts with various combinations of  
24 prevalence or penetrance values.

## 1 **Results**

### 2 **BDCS is caused by small tandem intergenic duplications at chromosome Xq26.1**

3 Previous mapping has linked BDCS to an 11.4 Mb interval on chromosome Xq25-27.1  
4 [24, 25]. To further investigate the genetic etiology of BDCS, we ascertained eight  
5 families (F1-8) with individuals affected with BDCS (Fig. 1), including five previously  
6 published families (F1 [25], F4 [26], F5 [27], F7 [3], and F8 [28]). Whole exome (F1 and  
7 F2) or genome (F3) sequencing did not identify putative disease-causing variants within  
8 the Xq26.1 locus previously linked to BDCS [25]. Array-comparative genomic  
9 hybridization in at least one affected individual from F1, F2, F3, and F7 revealed small  
10 intergenic Xq26.1 gains of varying sizes (Fig. 2A) [25]. Further, qPCR assays in affected  
11 individuals from the other four families (F4, F5, F6, and F8) were consistent with gains at  
12 this locus and confirmed that the gains segregated with BDCS in the multiplex families or  
13 were *de novo* in the two simplex cases (Fig. 2B). Long-range gap-PCR to amplify the  
14 duplication junctions followed by Sanger sequencing defined the head-tail junctions for  
15 F2, F4, F6 and F7 and confirmed these gains as tandem duplications (Fig. 2C). The  
16 duplications in F1, F3, F5, and F8 could not be differentiated further, likely due to  
17 homologous L1 elements (Fig. 2D). In the three multiplex families (F1, F3 and F8) with  
18 seemingly identically sized duplications, SNP haplotype mapping proved their likely  
19 independent origin (Fig. 2E). No gains were identified by qPCR in 215 unrelated  
20 European controls (139 females and 76 males). The largest gain was detected in F2  
21 (135kb) and the 18kb gain in F6 defined the smallest shared overlapping region (hg19,  
22 chrX:130,341,750-130,360,310) (Fig. 2D).

23 No similar sized exclusively intergenic gains overlapping with the shared duplicated  
24 region defined by the eight families were detected in control populations (Fig. S1). One

1 entirely intergenic gain (nsv517789; chrX:130,234,700-130,340,623) of ~100kb, in an  
2 individual with no known phenotype was noted, but it did not overlap the shared  
3 duplicated region as defined by our study (Fig. S1) [29]. Another gain of >380kb  
4 (nsv528179; chrX:130,300,617-130,680,930) overlapping the shared BDCS duplicated  
5 region, in an individual with no known phenotype was noted, but it extended ~295kb  
6 beyond the most telomeric boundary of the BDCS associated duplications and  
7 encompasses *IGSF1* (Fig. S1). Other larger chromosome X duplications encompassing  
8 the region have also been reported in individuals without BDCS [29].

### 9 **The Xq26.1 duplications in BDCS likely dysregulate ARHGAP36**

10 None of the BDCS duplications encompass protein-coding genes, suggesting  
11 dysregulation of flanking genes as the likely disease-associated mechanism (Fig. 2D).  
12 BDCS is considered to be a disorder of the hair follicle [5]. We, therefore, performed  
13 immunofluorescence for the two flanking genes, ARHGAP36 and IGSF1 in control hair  
14 follicles. In anagen, IGSF1 was present in the terminally differentiated inner root sheath  
15 (IRS), but was absent from the actively proliferating hair matrix or bulge, the main  
16 reservoir of hair follicle epithelial stem cells (Fig. S3a). In telogen, there was no evidence  
17 of IGSF1 staining (data not shown). In contrast, ARHGAP36 was present in a small  
18 number of hair follicle cells in the outer root sheath (ORS) at the level of the stem cell  
19 bulge, in both anagen (Fig. 3A) and telogen (Fig. 3C). Notably, it is at the end of telogen  
20 that the stem cells located in the secondary hair germ and bulge regions are activated to  
21 resume hair growth [30]. There was no obvious difference regarding ARHGAP36 positive  
22 cell numbers in the ORS between hair follicles from healthy control and an individual with  
23 BDCS (F5:II-1) in anagen (Fig. 3A and 3B). In contrast, in telogen hair follicle from F5:II-1,  
24 there was a marked increase in the number of ARHGAP36 positive cells around the

1 epithelial stem cell compartment (Fig. 3D), comparing with healthy control in telogen  
2 (Fig.3C) and the same patient in anagen (Fig. 3B). Moreover, immunofluorescence in  
3 histologically confirmed p63 positive BCC [31] from F4:III-4 showed strong staining for  
4 ARHGAP36 in a proportion of cells (Fig. 4A-B). Notably, the BCCs did not  
5 immunofluoresce for IGSF1 (Fig. S4a). We also detected striking ARHGAP36 staining of  
6 a trichoepithelioma from the same individual (Fig. 4C).

7 Next, to explore if ARHGAP36 could be relevant to sporadic BCCs, we determined the  
8 presence of ARHGAP36 in superficial (n=10), nodular (n=10), and infiltrative (n=10)  
9 sporadic BCCs. Similar to the BCCs in BDCS, ARHGAP36 was present in a small  
10 proportion of cells from all examined tumors (Fig. 4D-F), but was absent from all  
11 surrounding tissue.

## 12 ***ACTRT1* loss-of-function variants are unlikely to cause BDCS**

13 Knudson's two-hit hypothesis explains the mechanism of most inherited cancer  
14 syndromes, in which tumors occur following somatic loss of the only functional allele in  
15 an individual with a germline loss-of-function variant in a tumor suppressor gene [32].  
16 Presence of a single X-chromosome in males and X-inactivation in females, make  
17 X-linked inherited cancer predisposition syndromes due to tumor gene suppression  
18 highly unlikely. Previously, Bal *et al* reported loss-of-function *ACTRT1* variants in BDCS  
19 [28]. They identified a frameshift *ACTRT1* NM\_138289.3: c.547dup (p.Met183Asnfs\*17)  
20 variant (rs771087307), in two families with BDCS [28]. Variants in conserved non-coding  
21 elements flanking *ACTRT1* were also proposed as pathogenic. Our modelling revealed  
22 that the predicted maximum tolerated minor allele frequency (MAF) for the  
23 p.Met183Asnfs\*17 variant in population control data to be  $\sim 10^4$  times higher than

1 expected and could be reconciled only if the previous prevalence and penetrance  
2 estimates were inaccurate by several orders of magnitude (Fig. 5). Furthermore, other  
3 putative loss-of-function variants in the single exon of *ACTRT1* have been reported in  
4 both males and females without BDCS (Tables S1-2). Sanger sequencing in at least one  
5 affected individual from F1-F7 did not identify any rare coding variants in *ACTRT1*. As  
6 expected, the p. Met183Asnfs\*17 variant was present in the affected individuals (II-1 and  
7 III-1) in F8, as reported in the original association paper [28]. Importantly,  
8 immunofluorescent staining showed that *ACTRT1* was absent from both the  
9 disease-relevant regions of the hair follicles and tumors from individuals with BDCS (Fig.  
10 S3b and S4b). These combined data provide evidence that *ACTRT1* loss-of-function  
11 variants are unlikely to cause BDCS.

## 12 **DISCUSSION**

13 We show that BDCS is caused by small intergenic tandem duplications at Xq26.1 that  
14 encompass a minimum shared overlapping region of ~18Kb  
15 (chrX:130,341,750-130,360,310) (Fig. 2). Identification of a BDCS duplication at this  
16 locus in a family previously reported to have a causal *ACTRT1* variant is indicative of the  
17 likely benign nature of *ACTRT1* variants in individuals with BDCS [28, 33]. Importantly,  
18 our data suggests that larger Xq26.1 duplications *encompassing* the flanking genes in  
19 addition to the minimum common shared region may not result in BDCS. Also, smaller  
20 'non-coding' duplications at Xq26.1 that do not overlap the shared minimum overlapping  
21 region may not cause BDCS. These observations reflect the complexity of the genetic  
22 diagnosis of BDCS and assigning pathogenicity to Xq26.1 duplications in diagnostic labs  
23 will require careful consideration of the size and location of the copy number gains at this



1 position. Future studies will be needed to validate our findings and to refine the minimum  
2 common overlapping region for BDCS duplications.

3 To the best of our knowledge, BDCS is the first example of an inherited cancer  
4 predisposition disease caused by germline CNVs that do not encompass any protein  
5 coding genes. To explore the effects of the intergenic BDCS duplications, we first turned  
6 our attention to the two flanking genes, as our preliminary Hi-C assay on dermal  
7 fibroblasts from health control and BDCS patients showed that the of BDCS duplications  
8 did not disrupt the topologically associated domain (TAD), which contained *IGSF1* and  
9 *ARHGAP36* (Fig. S5). *IGSF1*, the flanking telomeric gene, encodes member 1 of the  
10 immunoglobulin superfamily and loss-of-function variants in this gene cause an X-linked  
11 recessive syndrome of central hypothyroidism and testicular enlargement (MIM 300888)  
12 [34]. There is no known role of *IGSF1* in Hedgehog-Patched-Gli pathway. *ARHGAP36*,  
13 the flanking centromeric gene, encodes a Smoothened (Smo)-independent positive  
14 regulator of the sonic hedgehog (shh) pathway and its expression is upregulated in  
15 medulloblastomas, a Hedgehog-Patched-Gli pathway-related tumor [35, 36]. Variants in  
16 *ARHGAP36* have not been associated with any inherited disorder. Hence, the known  
17 biological role makes *ARHGAP36* an excellent candidate for BDCS pathology. In this  
18 context, our finding of *ARHGAP36* expression is increased in telogen phase in follicles,  
19 in BCCs and trichoepithelioma from BDCS patients (Fig. 3D, 4B and 4C) makes it highly  
20 likely that *ARHGAP36* dysregulation is the likely disease mechanism.

21 Interestingly, our results suggest that *ARHGAP36* is also relevant to sporadic BCC  
22 pathology (Fig. 4D-F). Of note, *ARHGAP36* did not co-localize with p63 (Fig. 4C-F) and  
23 the higher levels of *ARHGAP36* in trichoepithelioma than in the BDCS-associated and

1 sporadic BCCs suggests that dysregulation of ARHGAP36 may be an early step in the  
2 pathogenesis to BCC. ARHGAP36, therefore, could be an attractive therapeutic target  
3 for inhibition in individuals with both inherited and, the vastly more common,  
4 non-inherited forms of BCC.

5 ARHGAP36 is a member of the Rho GAP family of regulatory proteins, which deactivate  
6 Rho proteins. Rac1 (a RhoGTPase) is essential for hair follicle stem cell function [37, 38].

7 With pulldown assays, we demonstrated interaction between ARHGAP36 and RAC1 (Fig.  
8 S6). This preliminary finding potentially uncovers a previously unknown role of  
9 ARHGAP36 and may explain the non-cancerous phenotypes of BDCS, namely  
10 hypotrichosis. However, further confirmatory studies will be required to solve the mystery  
11 underlying *ARHGAP36* dysregulation and BDCS phenotypes.

## 12 **Conclusions**

13 In summary, we have shown that small intergenic non-coding tandem duplications at  
14 Xq26.1 encompassing chrX:130,341,750-130,360,310 (hg19) cause BDCS. This is the  
15 first example of an inherited cancer predisposition disease caused by germline  
16 non-coding CNVs. We propose that the duplications result in the dysregulation of  
17 *ARHGAP36* that underlies BDCS pathology. Our findings reconcile the molecular  
18 mechanism of BDCS, a tumor-predisposition syndrome, with its X-linked inheritance  
19 pattern. We also suggest that ARHGAP36 is relevant to sporadic BCCs and a potential  
20 therapeutic target.

## 21 **Funding:**

22 XZ is supported by the National Natural Science Foundation of China (NSFC)  
23 [grant number 81788101], the CAMS Innovation Fund for Medical Sciences

1 (CIFMS) [grant numbers 2021-I2M-1-018 and 2016-I2M-1-002] and National Key  
2 Research and Development Program of China [grant number 2016YFC0905100].  
3 MJS, DGE, RP and WGN are supported by the Manchester NIHR Biomedical  
4 Research Centre (IS-BRC-1215-20007). ZM supported by NHS Grampian  
5 Biorepository (NHS Scotland). JME is funded by a postdoctoral research  
6 fellowship from the Health Education England Genomics Education Programme  
7 (HEE GEP).

#### 8 **Authors' contributions:**

9 XZ, JF and WGN conceived the project. XZ designed and coordinated the genetics study.  
10 WGN and SB designed and directed the histopathology study. YL, YH, MGK, MJB, GO,  
11 SE, JTL and RP conducted the experiments. YL, YH, JHS, DP, GMB, MGK, CF, CS, SC,  
12 JME, MJS, IHC, SB, AT, MvG, AB, JF, ZM, FA, AR, DGE, LJMTW-P, MAMvS, DCM, JF  
13 and WGN collected clinical data. SB, YL, YH, JF, WGN and XZ wrote the manuscript.

#### 14 **Availability of data and materials**

15 WES, WGS and array-comparative genomic hybridization datasets have not been  
16 deposited in a public repository due to privacy and ethical restrictions, but are available  
17 from the corresponding authors on reasonable request.

#### 18 **Ethics approval and consent to participate**

19 The procedures followed were in accordance with the ethical standards of the  
20 responsible committee on human experimentation (institutional and national), and written  
21 informed consents to take part in the present study, as well as to have the results of this  
22 work published were obtained from the participants .

#### 23 **Competing interests**

1 The authors declare no competing interests.

## 2 **References:**

- 3 1. Bazex AD, A.; Christol, B. Genodermatose complexe de type indetermine associant une  
4 hypotrichose, un etat atrophodermique generalise et des degenerescences cutanees multiples.  
5 *Bull Soc Franc Derm Syph.* 1964;71
- 6 2. Plosila M, et al. The Bazex syndrome: follicular atrophoderma with multiple basal cell  
7 carcinomas, hypotrichosis and hypohidrosis. *Clin Exp Dermatol.* Jan 1981;6(1):31-41.  
8 doi:10.1111/j.1365-2230.1981.tb02265.x
- 9 3. Herges A, et al. [Bazex-Dupre-Christol syndrome. Follicular atrophoderma, multiple basal cell  
10 carcinomas and hypotrichosis]. *Hautarzt.* Jun 1993;44(6):385-91. Das  
11 Bazex-Dupre-Christol-Syndrom. Follikulare Atrophodermie, multiple Basaliome und  
12 Hypotrichose.
- 13 4. Vabres P, de Prost Y. Bazex-Dupre-Christol syndrome: a possible diagnosis for basal cell  
14 carcinomas, coarse sparse hair, and milia. *Am J Med Genet.* Mar 15 1993;45(6):786.  
15 doi:10.1002/ajmg.1320450628
- 16 5. Goeteyn M, et al. The Bazex-Dupre-Christol syndrome. *Arch Dermatol.* Mar  
17 1994;130(3):337-42.
- 18 6. Yung A, Newton-Bishop JA. A case of Bazex-Dupre-Christol syndrome associated with  
19 multiple genital trichoepitheliomas. *Br J Dermatol.* Sep 2005;153(3):682-4.  
20 doi:10.1111/j.1365-2133.2005.06819.x
- 21 7. Barcelos AC, Nico MM. Bazex-Dupre-Christol syndrome in a 1-year-old boy and his mother.  
22 *Pediatr Dermatol.* Jan-Feb 2008;25(1):112-3. doi:10.1111/j.1525-1470.2007.00596.x
- 23 8. Castori M, et al. Bazex-Dupre-Christol syndrome: an ectodermal dysplasia with skin  
24 appendage neoplasms. *Eur J Med Genet.* Jul-Aug 2009;52(4):250-5.  
25 doi:10.1016/j.ejmg.2008.12.003
- 26 9. Kallam AR, et al. Basal Cell Carcinoma Developing from Trichoepithelioma: Review of  
27 Three Cases. *J Clin Diagn Res.* Mar 2016;10(3):PD17-9. doi:10.7860/JCDR/2016/15432.7464
- 28 10. Hahn H, et al. Mutations of the human homolog of Drosophila patched in the nevoid basal  
29 cell carcinoma syndrome. *Cell.* Jun 14 1996;85(6):841-51. doi:10.1016/s0092-8674(00)81268-4
- 30 11. Johnson RL, et al. Human homolog of patched, a candidate gene for the basal cell nevus  
31 syndrome. *Science.* Jun 14 1996;272(5268):1668-71. doi:10.1126/science.272.5268.1668
- 32 12. Smith MJ, et al. Germline mutations in SUFU cause Gorlin syndrome-associated childhood  
33 medulloblastoma and redefine the risk associated with PTCH1 mutations. *J Clin Oncol.* Dec 20  
34 2014;32(36):4155-61. doi:10.1200/JCO.2014.58.2569
- 35 13. Peris K, et al. Diagnosis and treatment of basal cell carcinoma: European consensus-based  
36 interdisciplinary guidelines. *Eur J Cancer.* Sep 2019;118:10-34. doi:10.1016/j.ejca.2019.06.003
- 37 14. Epstein EH. Basal cell carcinomas: attack of the hedgehog. *Nat Rev Cancer.* Oct  
38 2008;8(10):743-54. doi:10.1038/nrc2503
- 39 15. Peterson SC, et al. Basal cell carcinoma preferentially arises from stem cells within hair  
40 follicle and mechanosensory niches. *Cell Stem Cell.* Apr 2 2015;16(4):400-12.  
41 doi:10.1016/j.stem.2015.02.006
- 42 16. Youssef KK, et al. Adult interfollicular tumour-initiating cells are reprogrammed into an  
43 embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. *Nat Cell Biol.*  
44 Dec 2012;14(12):1282-94. doi:10.1038/ncb2628

- 1 17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
- 2 *Bioinformatics*. Jul 15 2009;25(14):1754-60. doi:10.1093/bioinformatics/btp324
- 3 18. McKenna A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing
- 4 next-generation DNA sequencing data. *Genome Res*. Sep 2010;20(9):1297-303.
- 5 doi:10.1101/gr.107524.110
- 6 19. Drmanac R, et al. Human genome sequencing using unchained base reads on self-assembling
- 7 DNA nanoarrays. *Science*. Jan 1 2010;327(5961):78-81. doi:10.1126/science.1181498
- 8 20. Carnevali P, et al. Computational techniques for human genome resequencing using mated
- 9 gapped reads. *J Comput Biol*. Mar 2012;19(3):279-92. doi:10.1089/cmb.2011.0201
- 10 21. Banka S, et al. Identification and characterization of an inborn error of metabolism caused by
- 11 dihydrofolate reductase deficiency. *Am J Hum Genet*. Feb 11 2011;88(2):216-25.
- 12 doi:10.1016/j.ajhg.2011.01.004
- 13 22. Sun M, et al. Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are
- 14 caused by genomic duplications involving the long range, limb-specific SHH enhancer. *J Med*
- 15 *Genet*. Sep 2008;45(9):589-95. doi:10.1136/jmg.2008.057646
- 16 23. Whiffin N, et al. Using high-resolution variant frequencies to empower clinical genome
- 17 interpretation. *Genet Med*. Oct 2017;19(10):1151-1158. doi:10.1038/gim.2017.26
- 18 24. Vabres P, et al. The gene for Bazex-Dupre-Christol syndrome maps to chromosome Xq. *J*
- 19 *Invest Dermatol*. Jul 1995;105(1):87-91. doi:10.1111/1523-1747.ep12313359
- 20 25. Parren LJ, et al. Linkage refinement of Bazex-Dupre-Christol syndrome to an 11.4-Mb
- 21 interval on chromosome Xq25-27.1. *Br J Dermatol*. Jul 2011;165(1):201-3.
- 22 doi:10.1111/j.1365-2133.2011.10219.x
- 23 26. Kidd A, et al. A Scottish family with Bazex-Dupre-Christol syndrome: follicular
- 24 atrophoderma, congenital hypotrichosis, and basal cell carcinoma. *J Med Genet*. Jun
- 25 1996;33(6):493-7. doi:10.1136/jmg.33.6.493
- 26 27. Torrelo A, et al. What syndrome is this? Bazex-Dupre-Christol syndrome. *Pediatr Dermatol*.
- 27 May-Jun 2006;23(3):286-90. doi:10.1111/j.1525-1470.2006.00237.x
- 28 28. Bal E, et al. Mutations in ACTRT1 and its enhancer RNA elements lead to aberrant
- 29 activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas. *Nat Med*. Oct
- 30 2017;23(10):1226-1233. doi:10.1038/nm.4368
- 31 29. MacDonald JR, et al. The Database of Genomic Variants: a curated collection of structural
- 32 variation in the human genome. *Nucleic Acids Res*. Jan 2014;42(Database issue):D986-92.
- 33 doi:10.1093/nar/gkt958
- 34 30. Schneider MR, et al. The hair follicle as a dynamic miniorgan. *Curr Biol*. Feb 10
- 35 2009;19(3):R132-42. doi:10.1016/j.cub.2008.12.005
- 36 31. Reis-Filho JS, et al. p63 expression in normal skin and usual cutaneous carcinomas. *J Cutan*
- 37 *Pathol*. Oct 2002;29(9):517-23. doi:10.1034/j.1600-0560.2002.290902.x
- 38 32. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci*
- 39 *U S A*. Apr 1971;68(4):820-3. doi:10.1073/pnas.68.4.820
- 40 33. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: a joint
- 41 consensus recommendation of the American College of Medical Genetics and Genomics and the
- 42 Association for Molecular Pathology. *Genet Med*. May 2015;17(5):405-24.
- 43 doi:10.1038/gim.2015.30
- 44 34. Sun Y, et al. Loss-of-function mutations in IGSF1 cause an X-linked syndrome of central
- 45 hypothyroidism and testicular enlargement. *Nat Genet*. Dec 2012;44(12):1375-81.
- 46 doi:10.1038/ng.2453
- 47 35. Rack PG, et al. Arhgap36-dependent activation of Gli transcription factors. *Proc Natl Acad*
- 48 *Sci U S A*. Jul 29 2014;111(30):11061-6. doi:10.1073/pnas.1322362111

- 1 36. Beckmann PJ, et al. Sleeping Beauty Insertional Mutagenesis Reveals Important Genetic  
2 Drivers of Central Nervous System Embryonal Tumors. *Cancer Res.* Mar 1 2019;79(5):905-917.  
3 doi:10.1158/0008-5472.CAN-18-1261  
4 37. Reijnders MRF, et al. RAC1 Missense Mutations in Developmental Disorders with Diverse  
5 Phenotypes. *Am J Hum Genet.* Sep 7 2017;101(3):466-477. doi:10.1016/j.ajhg.2017.08.007  
6 38. Behrendt K, et al. A function for Rac1 in the terminal differentiation and pigmentation of  
7 hair. *J Cell Sci.* Feb 15 2012;125(Pt 4):896-905. doi:10.1242/jcs.091868  
8

## 9 **Figure Legends**

10 **Fig. 1. Individuals with BDCS and pedigrees of families.** Pedigrees of all families  
11 included in this study are shown. Standard symbols have been used to denote sex and  
12 affected status. \* indicates individuals from whom DNA samples were available for  
13 analysis. At least one affected individual from each family underwent *ACTRT1* Sanger  
14 sequencing including its full coding region and the 5' and 3' untranslated regions. Figures  
15 of BDCS patients from current study were available from the corresponding authors on  
16 reasonable request.

17 **Fig. 2. Small intergenic duplications of chromosome Xq26.1 in individuals with**  
18 **BDCS.** (A) Array comparative genomic hybridisation in four individuals with BDCS,  
19 demonstrating intergenic copy number gains at the Xq26.1. The top four panels show the  
20 corresponding copy number of F1:IV-3, F2:II-2, F3:III-3 and F7:II-1 between chrX:  
21 130,150,000-130,450,000 (hg19), respectively. The bottom panel shows the flanking  
22 RefSeq protein coding genes. (B) qPCR of an amplicon at the duplicated locus  
23 demonstrating normal copy number (white bar) in the unaffected mothers of individuals  
24 F5:II-1 and F6:II-1 consistent with de novo origin of the duplications in these affected  
25 (black bar) individuals. In other families the presence of the duplications segregated with  
26 the phenotype. \*indicates different primer pair was used in F7 compared with other  
27 families. (C) The specific breakpoints for the duplications defined in affected individuals

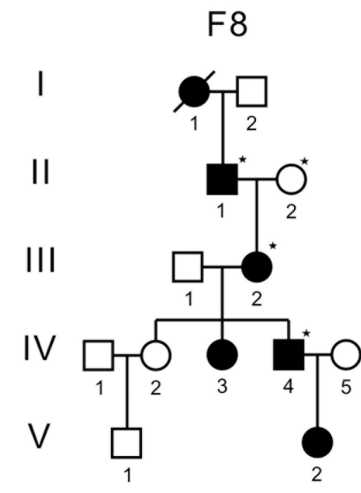
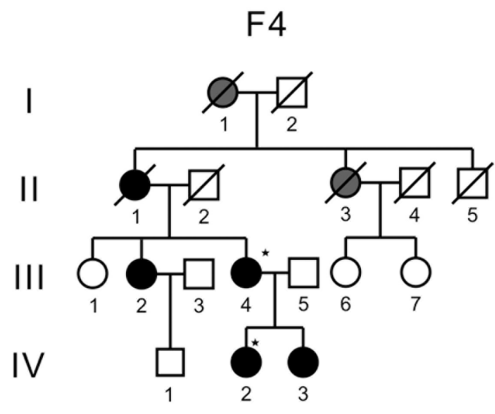
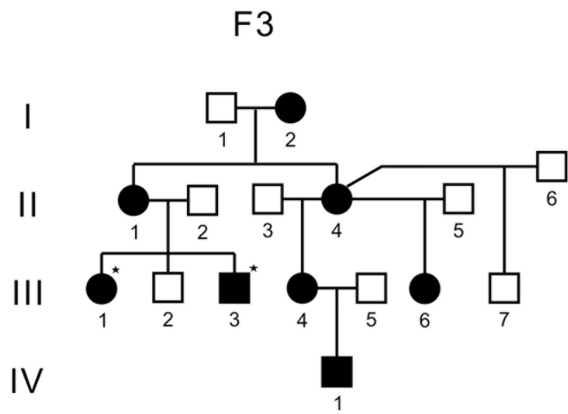
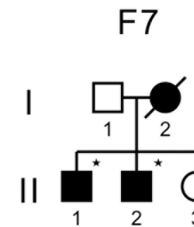
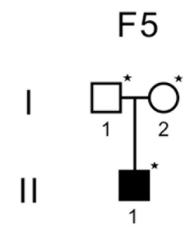
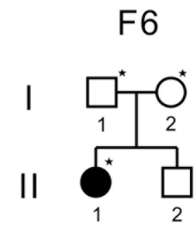
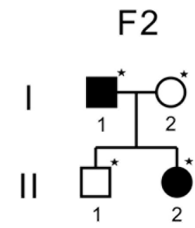
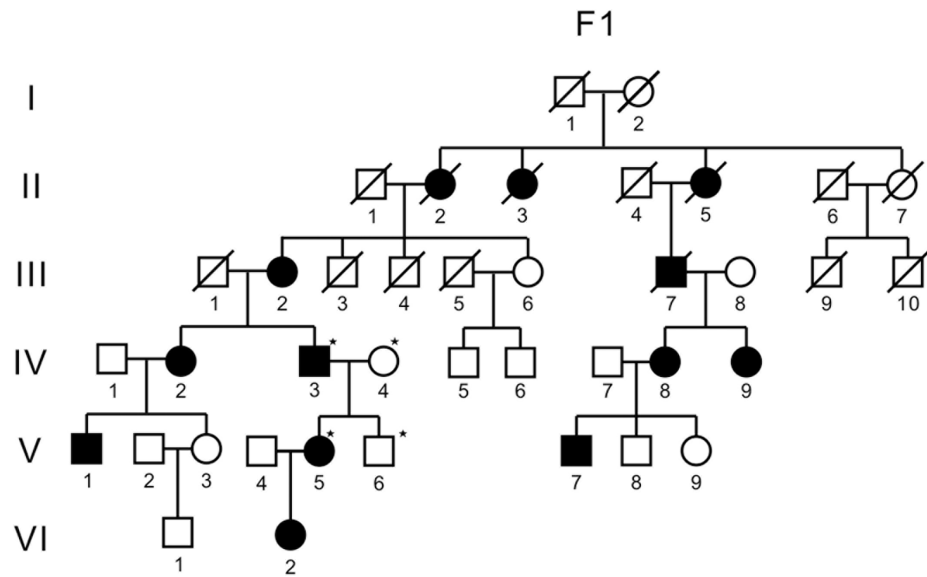
1 from families F2, F4, F6 and F7. (D) A cartoon of the duplications at Xq26.1-26.2 from  
2 the eight families with BDCS, with individual F6:II-1 defining the boundaries of the critical  
3 interval. (E) Long-range gap PCR in families' members of F1, F3, F5 and F8 had  
4 seemingly identically sized duplications (top). Haplotype analysis with four common  
5 single nucleotide polymorphisms (SNPs) (rs62603806, rs4240127, rs5932866,  
6 rs12559533) within the duplicated region demonstrates independent origins of F1, F3  
7 and F8 (down).

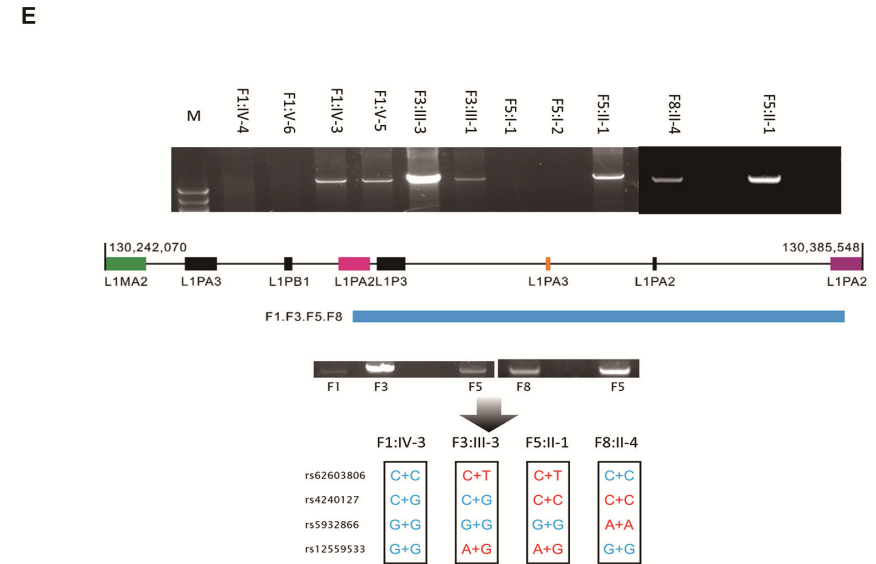
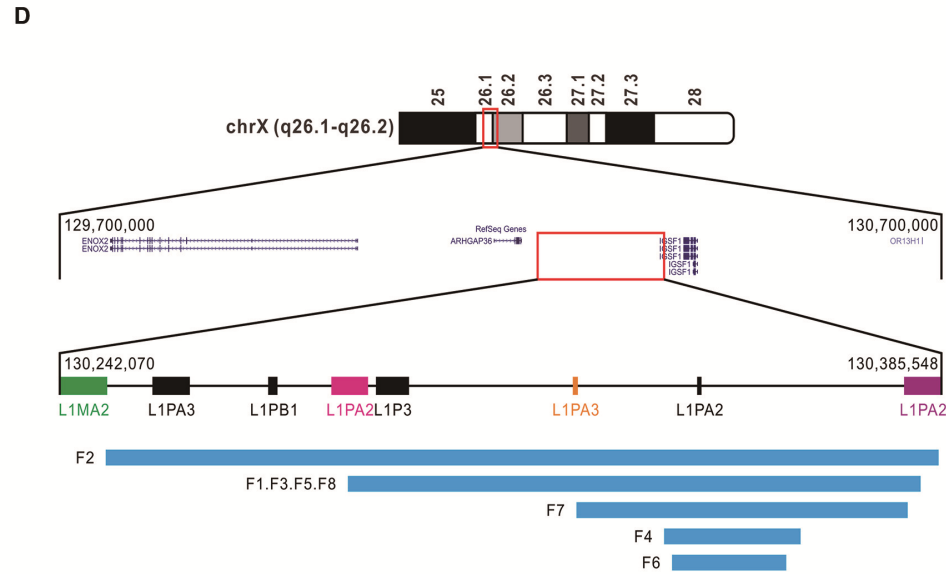
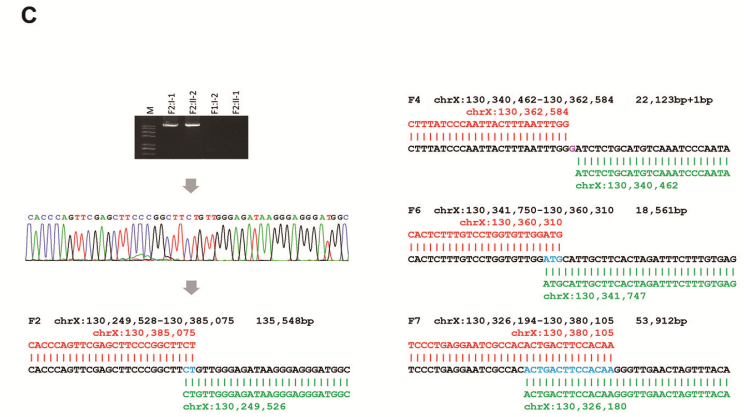
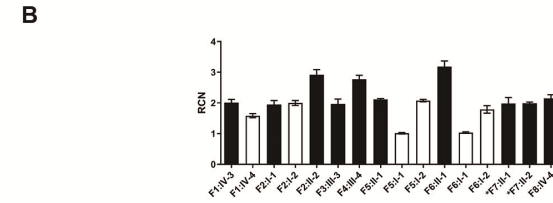
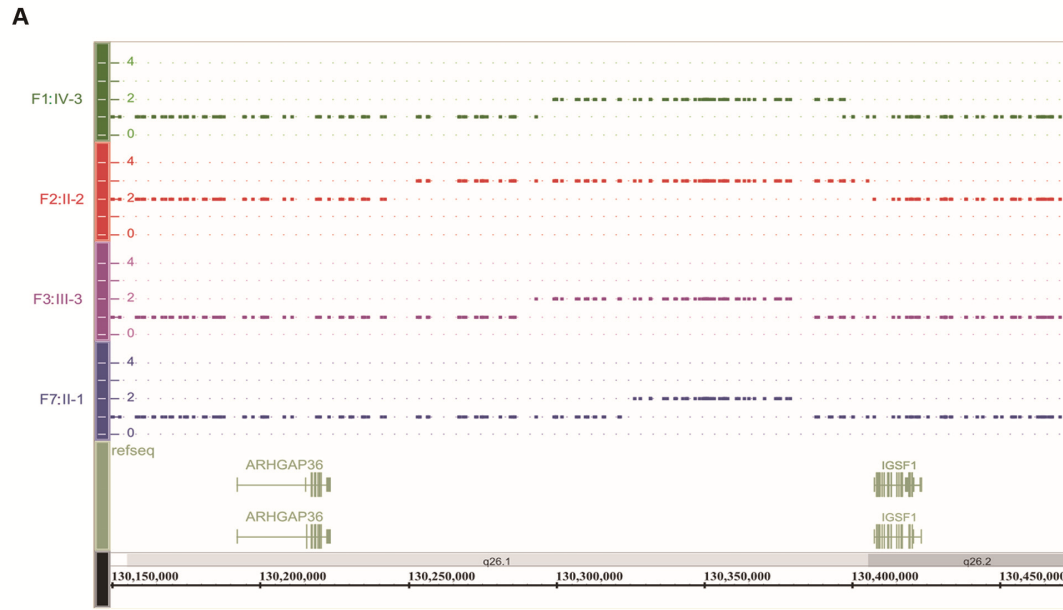
8 **Fig. 3. Immunofluorescence of hair follicles in anagen and telogen from normal**  
9 **healthy skin and an individual (F5:II-1) with BDCS for ARHGAP36.** A small number  
10 of cells in the stem cell bulge region stain positively (green) for ARHGAP36 in the hair  
11 follicle from normal healthy skin (A) and the skin from F5:II-1 (B). (C) A normal hair  
12 follicle in telogen stained for ARHGAP36 (green) demonstrating a small number of  
13 positively stained cells in the outer root sheath adjacent to K15 positive bulge stem cells  
14 (red). (D) A hair follicle in telogen from F5:II-1 showing an increased number of positively  
15 stained cells. Red stain for keratin 15. (DP = dermal papilla; ORS = outer root sheath; HS  
16 = hair shaft). Scale bar 50µm.

17 **Fig. 4. Immunofluorescence of ARHGAP36 in BDCS patient and sporadic BCCs.**  
18 (A-C) Pathological study of tumors. Haematoxylin and eosin staining (A) and  
19 immunofluorescence (B) in a BCC from an individual with BDCS (F4:III-4) demonstrating  
20 staining (green) for ARHGAP36 in the tumor but not in the surrounding tissue. Strong  
21 staining of a trichoepithelioma (C) for ARHGAP36 and P63 (pink) in the same individual.  
22 Superficial (D), nodular (E) and infiltrative (F) sporadic BCCs from individuals without  
23 BDCS stained for ARHGAP36 (green) and and P63 (pink). Scale bar 50µm.

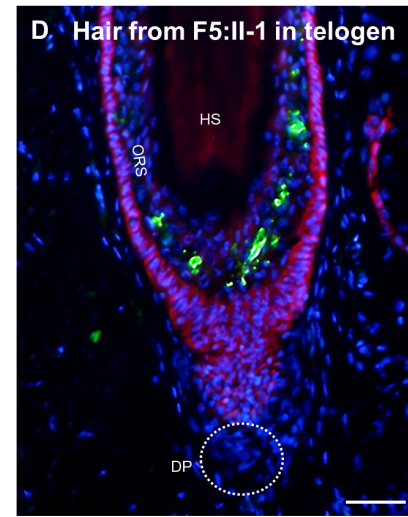
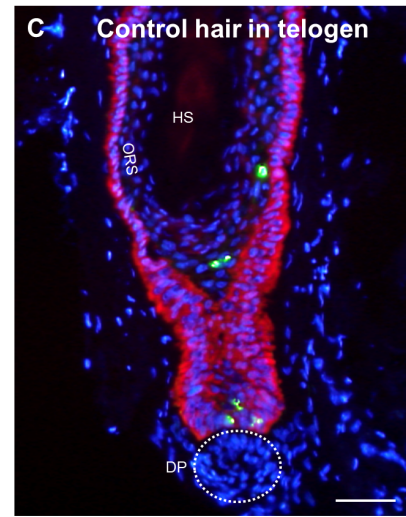
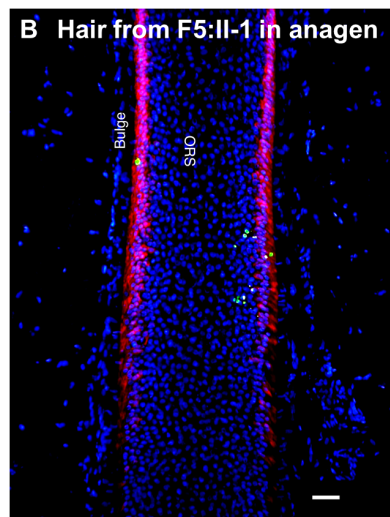
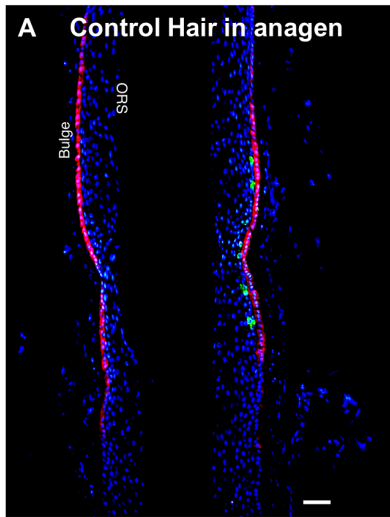
1 **Fig. 5. ACTRT1 loss-of-function variants are unlikely to cause BDCS.** Modelling of  
2 maximum tolerated allele counts (MTAC) for the *ACTRT1* NM\_138289.3:c.547dup  
3 (p.(Met183Asnfs\*17)) variant is shown. The left and right panels show MTACs against  
4 different levels of possible prevalence and penetrance respectively. In both panels, the  
5 blue lines model MTACs at the prevalence (left panel) or penetrance (right panel) levels  
6 consistent with the existing literature. The green lines model MTACs at much lower  
7 hypothetical constraints. Allele count estimates are for 204,684 alleles in the reference  
8 population at 0.95 confidence interval, 33% allelic heterogeneity and genetic  
9 heterogeneity of 1. Note that the observed population allele count (=373) of the variant,  
10 shown in broken red line, is significantly higher than estimated maximum tolerated allele  
11 counts across all the modelled scenarios.  
12

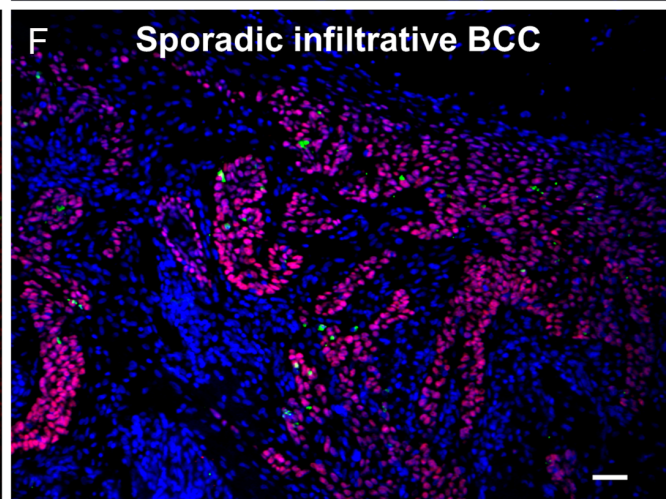
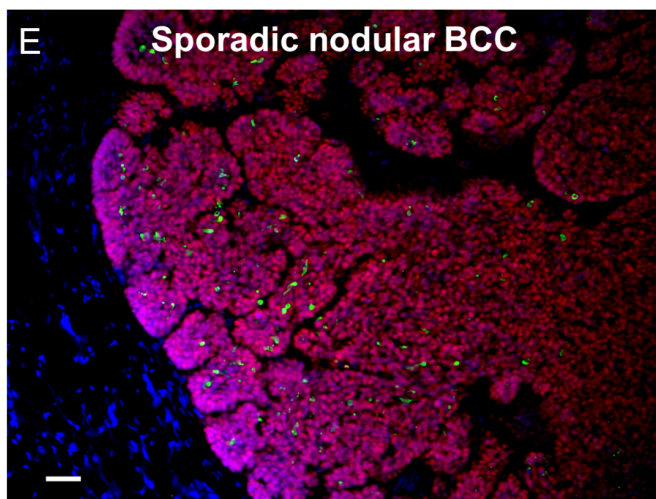
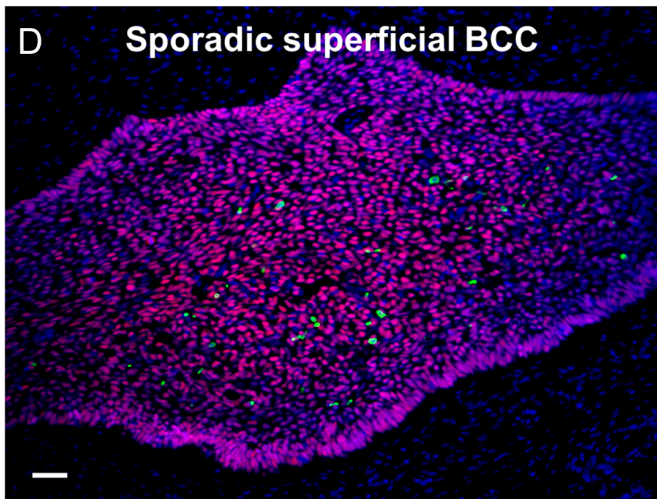
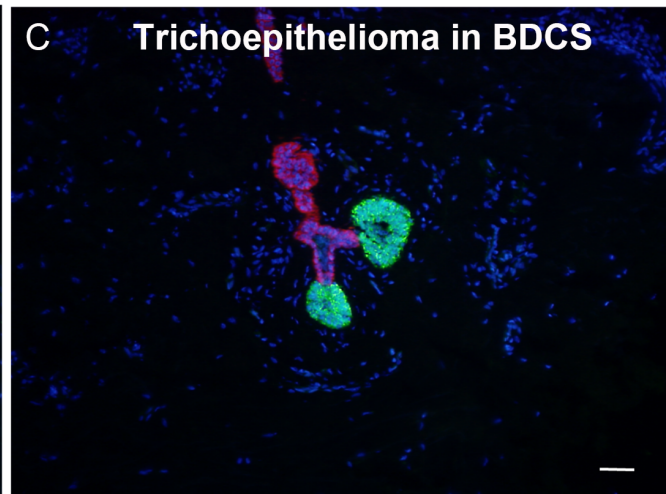
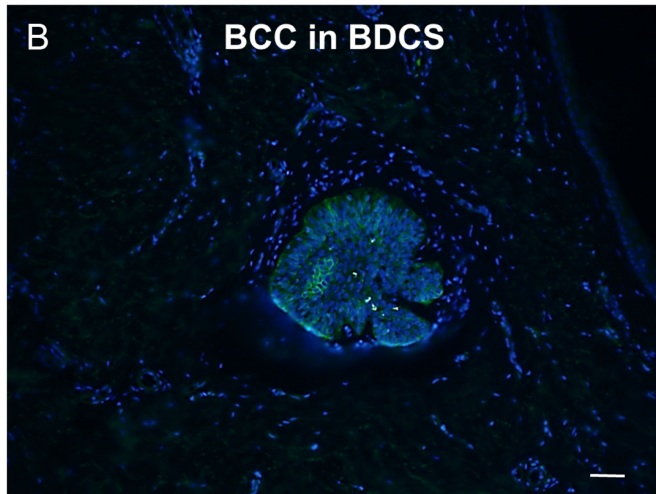
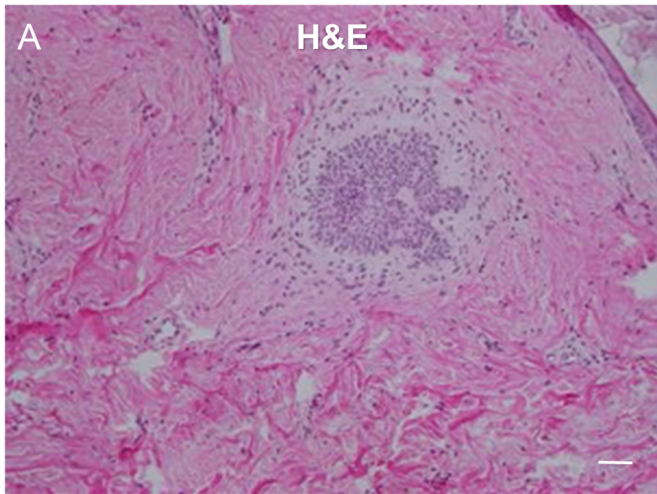


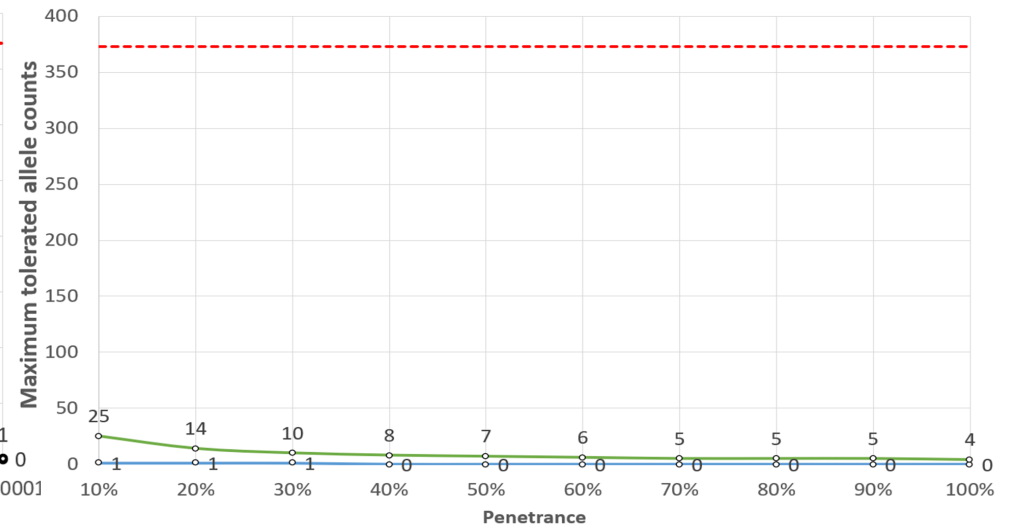
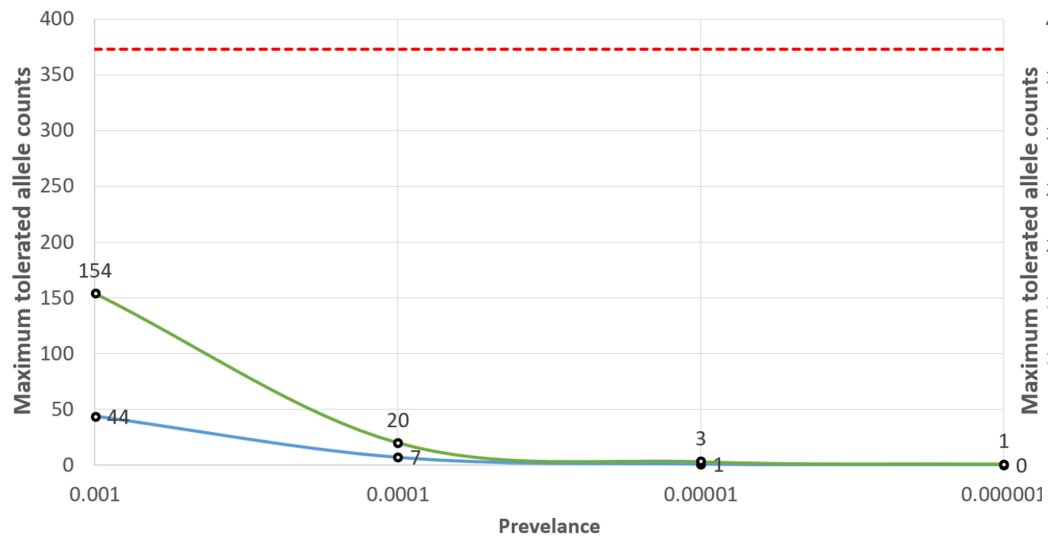




K15 ARHGAP36 DAPI







- At 100% penetrance (as expected from literature)
- At 25% penetrance
- - - Observed population allele count

- 0.000001
- 0.0001
- - - Observed population allele count