Phosphoproteomic profiling highlights CDC42 and CDK2 as key players in the regulation of the TGF-β pathway in *ALMS1* and *BBS1* knockout models

Brais Bea-Mascato[1,2], Girolamo Giudice[3], Iguaracy Pinheiro-de-Sousa[3], Carlos Solarat[1,2], Pablo Barbeito[1,2], Evangelia Petsalaki[3], Diana Valverde[1,2]

1 CINBIO, Universidad de Vigo, 36310 Vigo, Spain.

2 Grupo de Investigación en Enfermedades Raras y Medicina Pediátrica, Instituto de Investigación Sanitaria Galicia Sur (IIS Galicia Sur), SERGAS-UVIGO, Vigo, Spain

3 European Molecular Biology Laboratory-European Bioinformatics Institute, Wellcome Genome Campus, Hinxton CB10 1SD, UK

Correspondence: Diana Valverde

CINBIO Facultad de Biología, Universidad de Vigo, Campus As Lagoas-Marcosende s/n, 36310 Vigo, Spain

Tel +34 986 811 953

Email: dianaval@uvigo.es

List of supporting information:

Table S1. sgRNAs used to generate the KO model for the *BBS1* gene.

Table S2. Primers used for the characterisation of the mutations generated by the sgRNAs in the *BBS1* gene.

Figure S1. Characterisation of the *BBS1* KO model and quality control of phosphoproteomic datasets.

Figure S2. Cilia characterisation and ciliogenesis quantification in *BBS1* and *ALMS1* KO models.

The following supplementary information is available in the supplementary tables in the excel of this article, in separate tabs, and is therefore not included in this PDF:

Table S3. Results of protein identification with the proteome discovery in Batch 1 (basal).

Table S4. Results of protein identification with the proteome discovery in Batch 2 (with TGF-β stimulation).

Table S5. Results of differential phosphorylation analysis in batch 1 (basal).

Table S6. Results of differential phosphorylation analysis in batch 2 (TGF-β stimulation).

Table S7. Network diffusion results with nodes sorted by heat value in the KO model for *BBS1*.

Table S8. Network diffusion results with nodes sorted by heat value in the KO model for *ALMS1*.

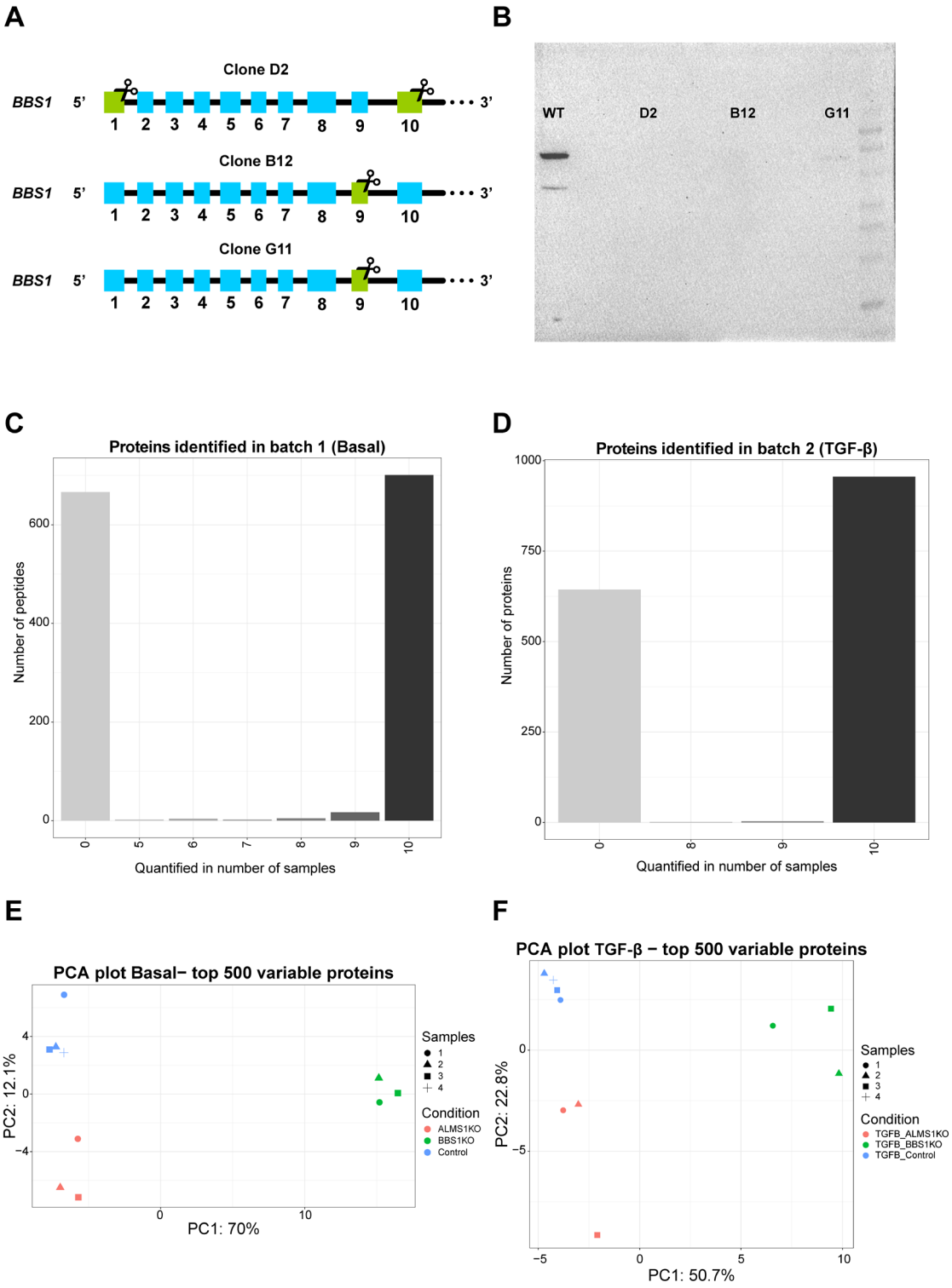
Supplementary Material:

Supplementary Table S1. sgRNAs used to generate the KO model for the *BBS1* gene. The letters in red signify the base pairs added for ligation of the oligos in the LentiCRISPRv2 plasmid. The sequence of the sgRNA is on the forward strand. The reverse strand sequence is shown for the generation of the double-stranded oligo to be inserted into the plasmid. In addition, the sgRNAs are shown with the off-target effect score and the on-target effect score.

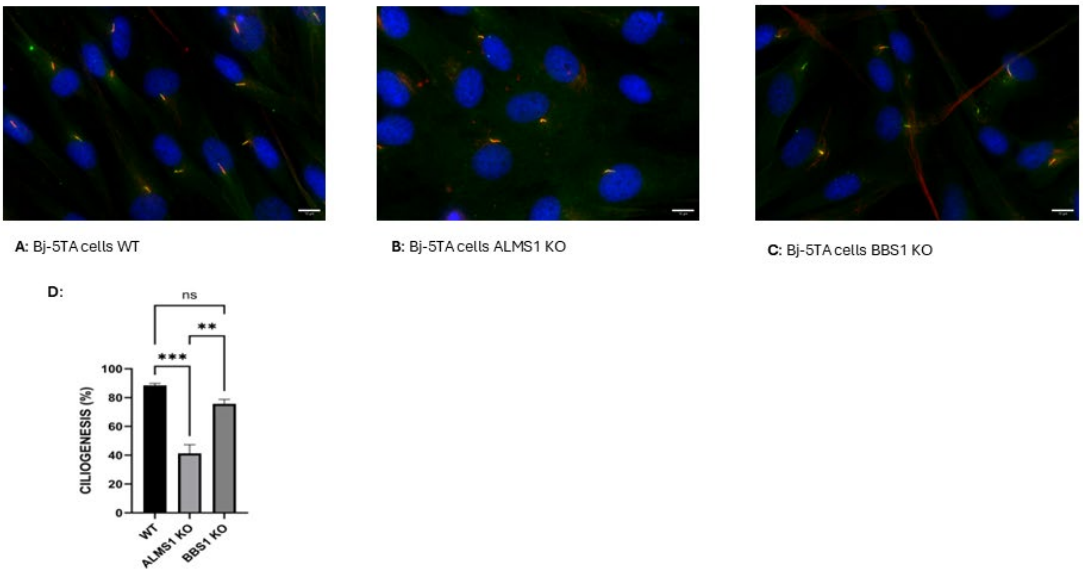
	ID	Sequence	On-target Score	Off-target Score	PAM	Strand
gRNA 1	gRNA 1.BBS1Exon 9. FWD	CACCGTGTTGAGTTCCGGCTTGCCG	61,7	92	CGG	+
	gRNA 1.BBS1Exon 9. RVS	aaacCGGCAAGCCGGAACTCAACAC				
gRNA 2	gRNA 2.BBS1Exon 1. FWD	CACCGCAGGCGTCGGAATCCGATG	69,2	92,4	AGG	-
	gRNA 2.BBS1Exon 1. RVS	aaacCATCGGATTCCGACGCCTGC				
gRNA 3	gRNA 3.BBS1Exon 10. FWD	CACCGTGTACCCGGATAAGTCCCAC	65	83	AGG	-
	gRNA 3.BBS1Exon 10. RVS	aaacGTGGGACTTATCCGGGTACAC				
gRNA 4	gRNA 4.BBS1Exon 4. FWD	CACCGGGCTTTCGGTCATCACCAG	71,3	75,3	TGG	-
	gRNA 4.BBS1Exon 4. RVS	aaacCTGGTGATGACCGAAAGCCC				

Supplementary Table S2. Primers used for the characterisation of the mutations generated by the sgRNAs in the *BBS1* gene. TM calculated with TM calculator (Thermo Fisher, Waltham, USA) and TM optimised empirically.

BBS1	Forward	Reverse	Amplified fragment (pb)	Tm calculated	Tm optimised
Exon 1	TGAGCCTGGGTGGGAAAG	GCCTCGGTTTCCCTATCT	297	59.8	60.5
Exon 4	AGGTCAGGCAGCTTGAAA	AGTGCCAGAGCTGGGGTC	300	58.5	63
Exon 9	TGGGACTTTAGACCAGGCAC	AAAGCCCACTCTCATCTTGC	293	56.9	59
Exon 10	AAGGTGGCAGAAGTGGAAAT	AGGGAACAGCCAGCGTCT	414	56.2	60.5


A
Clone D2
BBS1 5' 1 2 3 4 5 6 7 8 9 10 3'
Clone B12
BBS1 5' 1 2 3 4 5 6 7 8 9 10 3'
Clone G11
BBS1 5' 1 2 3 4 5 6 7 8 9 10 3'
B
WT
D2
B12
G11
C
Proteins identified in batch 1 (Basal)
Number of peptides
0
200
400
600
0
5
6
7
8
9
10
Quantified in number of samples
D
Proteins identified in batch 2 (TGF-β)
Number of proteins
0
250
500
750
1000
0
8
9
10
Quantified in number of samples
E
PCA plot Basal− top 500 variable proteins
PC2: 12.1%
4
0
−4
0
10
PC1: 70%
Samples
1
2
3
4
Condition
ALMS1KO
BBS1KO
Control
F
PCA plot TGF-β − top 500 variable proteins
PC2: 22.8%
0
−5
−5
0
5
10
PC1: 50.7%
Samples
1
2
3
4
Condition
TGFB_ALMS1KO
TGFB_BBS1KO
TGFB_Control

Supplementary Figure S1. Characterisation of the *BBS1* knockout model and quality control of phosphoproteomic datasets **(A)** Graphical representation of the mutated exon/s in the 3 knockout clones obtained after screening for the *BBS1* model. **(B)** Western-blot validation of the *BBS1* knockout clones (D2, B12, G11) **(C)** Number of proteins quantified in the different samples composing batch 1 (Basal) out of the total number of proteins identified **(D)** Number of proteins quantified in the different samples composing batch 2 (TGF-β) out of the total number of proteins identified **(E)** PCA plot of batch 1 (basal) with the top 500 quantified proteins with highest variation **(F)** PCA plot of batch 2 (TGF-β) with the top 500 quantified proteins with the highest variation.



Supplementary Figure S2. Cilia characterisation and ciliogenesis quantification in *BBS1* and *ALMS1* KO models. BJ-5ta cells were grown to confluence on coverslips, serum starved for 48 h to promote primary cilia formation, and fixed 10 min at RT in PBS+4% PFA. Representative immunofluorescence image of primary cilia in: (**A**) WT cells, (**B**) ALMS1 KO cells, (**C**) BBS1 KO cells. Primary cilia are stained with antibodies against ARL13B (green), acetylated α-tubulin (red) and DAPI-stained nuclei in blue. (**D**) Percentage of ciliated cells in control and KO cells (ALMS1 and BBS1) Data are mean ± SEM of n=3 independent experiments. Data were analyzed by one-way ANOVA with post- hoc Tukey multiple comparisons tests. Statistical significance is depicted as $p<0.002$(**), $p<0.001$(***), not significant (ns)