Human canonical CD157/Bst1 is an alternatively spliced isoform masking a previously unidentified primate-specific exon included in a novel transcript

Enza Ferrero, Nicola Lo Buono, Simona Morone, Rossella Parrotta, Cecilia Mancini, Alfredo Brusco, Alice Giacomino, Stefania Augeri, Antonio Rosal-Vela, Sonia García-Rodríguez, Mercedes Zubiaur, Jaime Sancho, Alessandra Fiorio Pla & Ada Funaro

CD157/Bst1 is a dual-function receptor and β-NAD⁺-metabolizing ectoenzyme of the ADP-ribosyl cyclase family. Expressed in human peripheral blood neutrophils and monocytes, CD157 interacts with extracellular matrix components and regulates leukocyte diapedesis via integrin-mediated signalling in inflammation. CD157 also regulates cell migration and is a marker of adverse prognosis in epithelial ovarian cancer and pleural mesothelioma. One form of CD157 is known to date: the canonical sequence of 318 aa from a 9-exon transcript encoded by BST1 on human chromosome 4. Here we describe a second BST1 transcript, consisting of 10 exons, in human neutrophils. This transcript includes an unreported exon, exon 1b, located between exons 1 and 2 of BST1. Inclusion of exon 1b in frame yields CD157-002, a novel proteoform of 333 aa: exclusion of exon 1b by alternative splicing generates canonical CD157, the dominant proteoform in neutrophils and other tissues analysed here. In comparative functional analyses, both proteoforms were indistinguishable in cell surface localization, specific mAb binding, and behaviour in cell adhesion and migration. However, NAD glycohydrolase activity was detected in canonical CD157 alone. Comparative phylogenetics indicate that exon 1b is a genomic innovation acquired during primate evolution, pointing to the importance of alternative splicing for CD157 function.
involved in regulation of MSC self-renewal, migration, and osteogenic differentiation\(^8\) but also highly expressed in the central nervous system\(^9\).

We previously showed that CD157 is a key player on orchestrating leukocyte and specific tumor cell type trafficking. CD157 modulates essential neutrophil and monocyte functions such as adhesion to the extracellular matrix (ECM)\(^2\), cell motility and transmigration\(^3\). Underlying these functional roles is the capacity of CD157 to bind with high affinity to components of the ECM such as fibronectin, fibrinogen, laminin-1 and type 1 collagen, within their heparin-binding domains\(^1\). CD157 binding confers outside-in signalling in leukocytes through structural and functional interactions with \(\beta_1\) and \(\beta_3\)-integrins\(^3\)\(^2\).

We also reported that CD157 was expressed in >90% of primary epithelial ovarian cancer (EOC)\(^1\)\(^4\) and in malignant pleural mesothelioma (MPM)\(^5\). In these specific cancer types, stemming from tissues with a common embryonic origin, high expression of CD157 correlates significantly with increased tumor aggressiveness and is an independent prognostic factor for overall survival in EOC and in biphasic MPM\(^6\). CD157 is widely expressed in M4 and M5 subtypes of acute myeloid leukemia (AML)\(^3\), notably in the patient group characterised by adverse prognosis according to the European LeukemiaNet (ELN) classification and CD157 is currently under investigation as a target for antibody-based immunotherapy in AML\(^7\)\(^8\). CD157 is also a promising marker for disease monitoring in B-cell precursor acute lymphoblastic leukemia (BCP-ALL) where high CD157 expression on malignant cells discriminates them from CD157-negative normal B cells from the same patient\(^9\).

While the receptor role of CD157 has been clearly delineated, the functional significance of CD157 as an enzyme is less clear at present. Human CD157 is reported to have ADP-ribosyl cyclase activity at acidic pH (4.0–6.5) and in the presence of metal (Zn\(^{2+}\) and Mg\(^{2+}\)) ions\(^1\)\(^0\)\(^1\)\(^2\), but its quantitative contribution to synthesizing the potent intracellular second messenger cyclic adenosine 5′-diphosphoribose (cADPR) by \(\beta\)-NAD\(^{2+}\) conversion is currently unknown. In contrast, the NAD\(^{4+}\)-glycohydrolase (EC 3.2.2.5) activity of CD157 is readily detectable\(^2\)\(^0\)\(^1\) and leads to release of adenosine 5′-diphosphoribose (ADPR) which can act both as a substrate and as a messenger\(^2\)\(^2\). In contrast to the observations made in human models, in mice there is evidence that CADPR generation by CD157 may be biologically relevant in mediating paracrine signalling in hematopoiesis\(^2\)\(^3\), in gut stem cell signalling in response to calorie restriction\(^2\)\(^4\) and in social behaviour\(^2\)\(^5\)\(^2\)\(^6\), envisaging a role for CD157 as a neuro-entero-immune regulator\(^2\). A further neuroregulatory role for human CD157 is under investigation following genome-wide association studies that have linked BST1 polymorphisms to susceptibility for Parkinson’s disease\(^2\)\(^7\)\(^2\)\(^8\). Parkinson-like behavioural abnormalities have been detected in bst1/bp3 knockout mice which show signs of non-motor dysfunction such as anxiety and depression-like symptoms\(^2\)\(^9\).

The human BST1 gene and cDNA were characterised in the 1990s using standard molecular cloning procedures\(^4\)\(^0\)\(^0\) with BST1 described as a 9-exon gene like its murine counterpart bst1/bp3\(^1\). BST1 maps to chromosome 4p15, next to its paralogue or gene duplicate CD38\(^2\)\(^9\). BST1 and CD38 ARC genes show conservation in their gene structure and protein sequence\(^3\). All the current functions of CD157 have been attributed to the only form known to date, the canonical protein of 318 amino acids believed to be derived by constitutive splicing of the 9-exon BST1 gene. However, in preparation for analysis of human BST1 transcript expression in normal human tissues by PCR, a primer pair designed to amplify a target sequence spanning exons 1 and 2 unexpectedly yielded two bands instead of one. As we are not aware of other forms of CD157 beyond the canonical, we investigated the PCR result which led to an upheaval in BST1/CST1 proteogenomics and we report here that: (i) by finding of a previously undescribed exon, exon 1b, we redefine BST1 as a 10-exon gene; (ii) there is a second BST1 transcript expressed in neutrophils and other tissues which encodes a novel and functionally divergent proteoform of 333 amino acids, CD157-002; (iii) production of canonical CD157, or CD157-001, requires exclusion of exon 1b by alternative splicing; and (iv) that exon 1b is a dynamically evolving exon present in phylogeny in BST1 orthologous genes in primates.

**Results**

**THP1 monocytic cell line and PMN express a BST1 transcript with a novel exon.** In preparation for analysis of the distribution of BST1 transcripts in human tissues by RT-PCR, an oligonucleotide primer pair was selected based on the orthodox 9-exon structure of the human BST1 gene, designed to amplify a 184 bp fragment spanning exons 1 and 2 of the canonical Bst1 transcript (BST1-001; Ensembl ENST00000265016; GenBank NM_004334.2) (Fig. 1a). THP1, a CD157-positive human monocytic leukaemia cell line, was selected as test template. Alongside the expected band, an additional faint upper band >200 bp long was observed (see the THP1 sample, Fig. 1b). To exclude a cell line artifact, the primer pair was tested on cDNA samples obtained from freshly isolated polymorphonuclear leukocytes (PMN) from 10 healthy adults, as PMN also show high expression of CD157. By PCR, all PMN samples showed the strong lower band/faint upper band couplet (Fig. 1b).

After agarose gel electrophoresis, the upper band was isolated and its sequence obtained, revealing it to consist of 229 bp. Although the sequence contained the expected 184 bp fragment, this was interrupted at the exon 1/exon 2 junction by inclusion of a 45 bp insert (Fig. 1c). As this 45 bp sequence appeared to represent a previously unrecognized exon, the Ensembl human database (GRCh38.p7) was analysed and a second BST1 transcript (BST1-002; ENST00000382346) was indeed found which included an exon formed by the same 45 bp sequence. This previously uncloned exon was designated exon 1b to maintain the numbering of homologous exons 2–9 shared with other vertebrate BST1 genes. Exon 1b is located within the BST1 gene on human chromosome 4p15, between exon 1 (1,587 bp downstream of the 3′ end) and exon 2 (550 bp upstream of the 5′ end) (Fig. 1c,d; Supplementary Figure S1), explaining the results obtained by PCR (Fig. 1c). The 45 bp sequence has all the hallmarks of an exon\(^3\)\(^3\): it is flanked by canonical AG/GT splice sites, shows a putative branch point (5′) 24 bp upstream of the 5′ end) (Fig. 1c,d; Supplementary Figure S1). The structure of human BST1 has been revised to that of a 10-exon gene (Fig. 1c).

Using either exon 1- or exon 1b-specific forward primers with a common exon 9 reverse primer, RT-PCR confirmed that THP1 and PMN have two full-length BST1 isoform transcripts which correspond to the two
Figure 1. Identification and characterization of human BST1 exon 1b. (a) Canonical 9-exon version of human BST1 showing the 184 bp amplicon expected by RT-PCR priming on exon 1 (arrow, forward primer) and exon 2 (arrow, reverse primer). (b) Agarose gel showing that THP1 cells and PMN from ten healthy donors yield doublet transcripts by RT-PCR. (c) Revised human BST1 gene structure with 10 exons following inclusion of a novel exon, 1b, identified experimentally by RT-PCR. The canonical transcript BST1-001 is alternatively spliced, skips exon 1b, and yields the 184 bp amplicon by RT-PCR; transcript BST1-002 is constitutively spliced, includes exon 1b, and yields the 229 bp amplicon by RT-PCR. (d) Schematic representation of the BST1 genomic locus on Homo sapiens chromosome 4 showing intron/exon boundaries of exon 1, 1b and 2. The orange box below exon 1b shows the 15 aa sequence added in-frame to the CD157 polypeptide by exon 1b inclusion. Consensus GT/AG splice sites are depicted in red. (e,f) Schematic models of the CD157 isoform polypeptides showing location of signal peptide (SP), ribosyl hydrolase domain (RHD) and GPI-anchoring propeptide in canonical CD157-001 (318 aa) and CD157-002 (333 aa) proteins, the latter including the exon 1b-encoded insert (orange band). (g,h) Results of comparative analysis of CD157 polypeptides by IUPRED which predicts intrinsic disorder. A tract of notable difference between CD157 proteofroms is boxed, corresponding to the region where exon 1b is included in CD157-002. (i) View of molecular models of CD157 dimer crystal structure PDB 1ISI (NCBI Structure): in background, CD157-001 dimer; in front, SWISS-MODEL-predicted dimer of CD157-002 with exon 1b-encoded peptides (both monomers) enclosed in yellow circles. (j) Alignment of the RHD from canonical CD157-001 and CD157-002. Secondary structures (alpha helices, beta sheets and turns) are indicated. Numbers in green show cysteine residues involved in disulphide bridge formation. RHD-002 sequence highlighted in yellow shows the 15 aa insert encoded by exon 1b, and peptide boxed in blue was identified in the proteomics database www.proteomicsdb.org.
transcripts described in the Ensembl database. These transcripts are: (i) alternatively spliced BST1-001 that skips exon 1b and encodes the canonical form of CD157 with 318 amino acids (Fig. 1e) and (ii) constitutively spliced BST1-002 (GenBank acc. no. 1747489) in which exon 1b is inserted in-frame to give an open reading frame encoding a predicted polypeptide of 333 aa, comprehensive of a signal peptide, complete ribosyl-hydrolase domain (Pfam 002267), conserved N-linked glycosylation sites and disulfide-forming cysteines, and C-terminal GPI pro-peptide (Fig. 1f). This proteoform is designated CD157-002. The 15 aa insert is located towards the N-term of the ribosyl-hydrolase domain (RHD). Analysis of the 318 and 333 aa variants by IUPRED, a software program that predicts ordered and disordered regions in proteins, suggests that the 15 aa insertion constitutes an intrinsically disordered region, while the remaining structured region of the RHD remains basically unchanged in CD157-002 with respect to canonical CD157 (CD157-001) (Fig. 1g,h). To date, only the 3D structure of canonical CD157 dimers is available; based on one of these structures (PDB) we have modelled the predicted CD157-002 3D structure (Fig. 1i), using the Swiss-model program. Exon 1b would be inserted after Arg63 in a 3_10 helix, i.e., a C-term extension of the α1 helix (Fig. 1j).

Finally, the RNA-based evidence was corroborated by peptide evidence indicating production of CD157-002 protein in vivo: a tryptic peptide (EPPVHTTQNK) encoded by exon 1b (aa 1–9 of the peptide) and exon 2 (aa 10–11) was identified in the www.proteomicsdb.org database as belonging to the 333 aa form of CD157 (acc. no. A6NC48). The peptide is shown boxed within the CD157-002 sequence in Fig. 1j.

**Comparative expression of BST1-001 and BST1-002 transcripts.** As was originally planned, BST1 transcript expression was studied by RT-PCR in a panel of human tissues, only now the experiments were carried out using the exon 1/exon 2 primer pair that could simultaneously detect BST-001 and -002 transcripts (Fig. 2a). BST-001 transcript was detected in all samples examined except for Raji Burkitt’s lymphoma cell line whereas the BST-002 transcript containing exon 1b was most evident in PMN, fetal and adult brain, cerebellum and colon. Using an exon 1b-specific forward primer, BST1-002 was easily detected not only in PMN but also in placenta, heart, lung, kidney, colon and spleen (Fig. 2a).

The Genotype-Tissue Expression (GTEx) portal offers a panoramic view of BST1 transcript expression in 53 nonpathological human tissues. According to this dataset, the BST1 gene is most highly transcribed in whole blood (presumably in myeloid cells, as erythroid and lymphoid cells are CD157-negative), with expression in other human tissues dwarfed in comparison (Fig. 2b). GTEx data was also available for selective expression of the canonical and BST1-002 isoforms, with BST1-001 the more expressed of the two (Fig. 2b, insets).

Therefore PMN were selected as the representative cell population to compare relative expression of BST1-001 and BST1-002 isoform transcripts by quantitative PCR (qPCR). Analysis of PMN RNA from 10 healthy adults showed remarkable inter-donor variability in the amount of both BST-001 and 002 transcripts (Supplementary Figure S2), and that the BST1-001 transcript was always more highly expressed than BST1-002, consistently representing ~75–80% of the sum of the two transcripts (Fig. 2c,d).

**Comparative functional analysis of CD157-001 and CD157-002 expressed in HeLa cells.** To begin the comparative analysis of the two CD157 proteoforms, their respective cDNAs were cloned into a pCDNA3-based expression vector and stable transfectants established in HeLa cells (whose phenotype is CD157-negative/CD38-negative). Cell surface expression was assessed by flow cytometry using the SY/11B5 mAb. This mAb recognized membrane-bound CD157-001 and CD157-002 with similar mean fluorescence intensity (Fig. 3a), indicating epitope conservation in CD157-002.

Comparative functional analysis of CD157-001 and CD157-002 expressed in HeLa cells. To begin the comparative analysis of the two CD157 proteoforms, their respective cDNAs were cloned into a pCDNA3-based expression vector and stable transfectants established in HeLa cells (whose phenotype is CD157-negative/CD38-negative). Cell surface expression was assessed by flow cytometry using the SY/11B5 mAb. This mAb recognized membrane-bound CD157-001 and CD157-002 with similar mean fluorescence intensity (Fig. 3a), indicating epitope conservation in CD157-002.

**CD157-002 is an efficient ECM receptor.** As canonical CD157 is a well-characterised receptor that mediates leukocyte, mesothelial and specific tumour (EOC, MPM) cell adhesion to ECM proteins, we conducted parallel studies to explore possible functional differences in the capacity of CD157-001 and CD157-002 to mediate cell attachment to fibronectin (FN), type I collagen (COL1) and vitronectin (VN) using the HeLa transfectant.
model. With respect to empty vector-transfected cells (designated as “mock”), expression of either CD157-001 or CD157-002 enhanced HeLa cell adhesion to microtiter plates coated with FN or COL1 (Fig. 4a). However, when comparing the two variants, HeLa/pBST1-002 bound FN and COL1 more efficiently than HeLa/pBST1-001, a difference that was statistically significant (p < 0.001). As previously described for canonical CD157, CD157-002

---

**Figure 2.** Expression profiles of BST1-001 and BST1-002 in human tissues. (a) RT-PCR analysis of BST1 expression in a multiple normal human tissue RNA panel. Amplified exons are indicated schematically on the right. β-actin is the internal control. A composite of representative gels is shown, within which each tissue type is indicated. Full-length gels are presented in Supplementary Figure 3. (b) Cross-tissue comparison of BST1 expression in 53 nonpathological human tissues from the GTEx resource. Expression is measured by RNA-seq and presented as Reads Per Kilobase transcript per Million reads (RPKM). The inserts show BST1 expression broken down into the 001 and 002 components. (c) Relative mRNA levels of BST1-001 and BST1-002 transcripts by RTqPCR expression analysis. The same cDNA amount of BST1-001 and BST1-002 was added to both assays. The threshold of GUSB was used to determine the ratio of each transcript. The relative expression of each transcript was calculated as the percentage of the sum of BST1-001 + BST1-002. Bars represent BST1-001 versus BST1-002 ratio, expressed as mean of three independent experiments. (d) Box plots showing expression of the two BST1 transcripts in PMN. Boxes indicate the range (25th–75th percentiles), whiskers indicate major and minor values, and the horizontal lines within the boxes indicate the median expression level of BST1-001 and BST1-002 transcripts. Results are expressed as relative BST1 mRNA variants expression in PMN from healthy subjects (****p < 0.0001, Student’s t test).
Figure 3. Comparative analyses of CD157 proteoforms expressed in HeLa cells and PMN. (a) Flow cytometric analysis of the expression of canonical CD157 (pBST1-001) and CD157-002 (pBST1-002) in HeLa cells. In each sample, $10^4$ cells were analysed; x-axis = fluorescence intensity, y-axis = number of cells (events). Shaded histograms indicate isotype-matched control IgG, while profiles represent expression of the indicated surface molecules in mock, pBST1-001 and pBST1-002 transfectants, respectively. (b) Western blot analysis of CD157-001 and -002 expressed in HeLa cells. Left panel, after SDS-PAGE in non-reducing conditions, detected with anti-CD157 mAb (SY/11b5); right panel, samples were treated with PNGase to remove N-linked oligosaccharides; full-length blots are presented in Supplementary Figure 4 (top panel). (c) Representative immunofluorescence images of the membrane localization of CD157-001 and CD157-002 proteins (in green) in HeLa cells transfected with expression vector of the two different BST1 isoforms. Nuclei were counterstained with propidium iodide (in blue). Cells were analysed using an Olympus FV300 laser scanning confocal microscope and were imaged using a 60X oil immersion objective (1.4 NA). Scale bar: 10 μm. (d) Western blot analysis of CD157-001 and -002 expressed in PMN (left panel) or U937 (right panel) after SDS-PAGE in reducing conditions, detected with polyclonal anti-CD157 IgG; where indicated, samples were treated with PNGase to remove N-linked oligosaccharides. One representative experiment is shown (n = 3). Full-length blot is presented in Supplementary Figure 5.
did not enhance adhesion to VN. Therefore, both CD157 proteoforms exhibit differential binding to discrete ECM proteins (i.e., to FN and COL1 but not to VN), with CD157-002 as efficient if not more so than CD157-001.

Wound healing assays were performed to compare the contribution of the two proteoforms to cell migration. In serum-free conditions, the migratory capacity of both CD157 proteoforms was similar, and greater than that of mock control (Fig. 4b). In the presence of 10% FCS, used as a source of ECM components, HeLa/CD157-002 again migrated more efficiently than cells expressing CD157-001 and this difference was statistically significant (p < 0.01).

We previously demonstrated that triggering CD157 by crosslinking with specific mAbs or adhesion to FN elicits phosphorylation of FAK (Tyr-397) leading to increased activation of the PI3K/AKT and MAPK/ERK1-2 pathways, and promotes actin cytoskeleton reorganization. Therefore, activation of FAK and its downstream signalling effectors AKT serine/threonine kinase and ERK1/2 were compared in HeLa cells expressing either CD157-001 or CD157-002 maintained in basal culture conditions without FCS or following the addition of FCS as a source of ECM components. Western blotting showed increased pFAK Tyr-397 phosphorylation in both CD157 transfectants compared to HeLa mock control both in serum-free conditions and in the presence of FCS.
of FCS. The levels of total FAK were comparable in all cell lines. AKT Ser-493 phosphorylation was upregulated in both CD157 transfectants in the presence of FCS. Phosphorylated ERK1/2 was notably increased in CD157-002-expressing cells, both with and without serum, with comparable amounts of total ERK2 in all cell lines. Taken together, these results show activation of the FAK motility pathway associated with the expression of either CD157-001 or -002, and increased ERK1/2 activation in HeLa/BST1-002 in line with the increased motility shown by these cells.

**Divergent enzymatic activity of canonical CD157 and CD157-002.** CD157 has generated much less interest as an enzyme than its relatives CD38 and alypia cyclase, possibly because CD157 appears to be a far less powerful catalyst than them. However, there are at least two experimental models (involving homeostasis of hematopoietic and intestinal stem cells), albeit in mouse, that are reliant upon CD157-derived cyclic ADP ribose, even though it may be produced in hormone-like (i.e., subnanomolar) concentrations. In the past, using sensitive HPLC methods, other labs demonstrated that human CD157 converted NAD⁺ to cyclic ADP ribose (minor product) and ADP ribose (major product), indicating the presence of ADP-ribosyl cyclase, NAD glycohydrolase and possibly cADPR hydrolase activities. Limited experiments with analog substrates have detected similar activities in human CD157, using etheno-NADE⁺ (conversion products not quantified) or NGD⁺ (ca. 100 nmole/ mg/min of cyclic GDPNAD produced).

Comparative analysis of CD157 enzymatic activity was carried out using whole HeLa cells expressing one or other of the two proteoforms. First, to test for NAD⁺ glycohydrolase activity, transfectants were incubated with the substrate 1,N(6)-etheno-NAD (ε-NAD⁺), a fluorescent analogue of β-NAD⁺. In the presence of canonical CD157, production of fluorescent etheno-ADP ribose (ε-ADPR) was observed, with the amount of product showing a linear increase over time (Fig. 4d). In contrast, when HeLa/CD157-002 cells were incubated with ε-NAD⁺, the fluorescence signal was negligible, indicating that CD157-002 does not share the NADase catalytic activity of canonical CD157. The NADase activity of the control Raji cell line, which strongly expresses CD38 but not CD157, was so high that after 10 min incubation ε-ADPR reached saturation and further measurement was not possible.

The nicotinamide guanine dinucleotide (NGD⁻) cyclization assay is frequently used as a proxy for ADP-ribosyl cyclase activity, with the caveat that these activities are not always equivalent. HeLa cell transfectants and controls were incubated with NGD⁻ to measure GDP-ribosyl cyclase activity via production of cyclic GDP-ribose (cGDPR, a fluorescent analogue of cADPR). This assay was almost negative with both canonical CD157 and CD157-002. In contrast, incubation of Raji cells with NGD⁻ yielded easily detectable cGDPR accumulation in cell supernatants, consistent with the ability of CD38 to convert NGD⁺ mostly to cyclic GDP ribose (cGDPR) and little to GDP (Fig. 4e). In summary, we detected a major difference in catalytic activity between CD157-001 and CD157-002.

**Conservation and dynamic evolution of BST1 exon 1b in primate phylogeny.** The 10-exon structure of BST1 in the human is a departure from the conserved 9-exon structure of the gene found in other mammalian (mouse, rat) and non-mammalian (chicken) species (EF, unpublished result). To establish if the 10-exon structure of the BST1 gene represents a human innovation, the evolutionary history of exon 1b was reconstructed by database BLAST analyses which retrieved homologous nucleotide sequences in 11 species, limited to primates (Fig. 5). In all of these species, the exon 1b-like sequence was found within the BST1 gene and always in orthologous positions between exons 1 and 2. Exon 1b-like sequences were not found in tarsiers (T. syrichta) nor Rodentia (mouse, rat), suggesting that this exon was acquired early in primate phylogeny after the separation of simians from tarsiers, estimated to have occurred approximately 67 million years ago (mya) (http://timetree.org).

Visual inspection of the alignment shows that although highly conserved, the exon 1b sequence has evolved over time, and apes have a shorter exon than monkeys because of an 11 bp deletion (Fig. 5). Among apes, whose genomes-to-proteomes databases are more complete than most of the monkey databases, there is evidence that not only human but also gorilla (Gorilla gorilla) and orangutans (Pongo Abelii, Pongo pygmaeus) express both constitutive 10-exon and alternatively spliced 9-exon transcripts. In gibbon (Nomascus leucogenys), chimpanzee (Pan troglodytes) and bonobo (Pan paniscus), we have currently found only 9-exon-derived transcripts and proteins. However, in chimp and bonobo, closer nucleotide sequence examination revealed that the consensus GT donor splice site is mutated to a nonconsensus AT donor splice sequence (Fig. 5). In gibbon, orangutans and gorilla, the GT at donor splice site has been replaced by an AT donor splice site, supporting the idea that the ancestral exon 1b splice site is the result of non conserved substitution (Fig. 5). This type of mutation is likely a lineage specific modification. The impact of exon 1b on CD157 in Old and New World monkeys is a work in progress as the relative databases are currently less informative, although there is some preliminary evidence in macaques of a different splicing pattern of exon 1b.

**Discussion**

The outcome of this study has gone beyond the expected results, leading us to: (i) redefine the structure of the human BST1 gene following the inclusion of a previously unidentified exon; (ii) identify a new BST1 transcript firstly in human neutrophils and then in other tissues, and demonstrate expression of both proteoforms in PMN; (iii) characterise the divergent functional characteristics of the novel proteofrm CD157-002 with respect to the canonical CD157-001 form; (iv) reveal that canonical CD157-001 is generated by alternative splicing; and (v) trace the origins of the new exon 1b within the primate phylogenetic tree.

Firstly, a static situation has changed. The human BST1 gene and cDNA were characterised in the 1990’s using standard molecular cloning procedures. The resulting 9-exon organization of the human gene was structurally homologous to the murine counterpart bst1/bp3. In addition, homologous cDNAs, cloned from mouse and rat,
were available so there was no reason to suspect that the human cDNA was anything other than the product of constitutive splicing whereby all of a gene's exons are joined together in the order in which they occur in the genome. This static image of BST1 has now started to change, thanks to the chance discovery of a novel BST1 exon found by experiment, and supported by abundant genomic, transcriptomic and proteomic data available in silico.

The process of joining different combinations of exons by AS is a powerful means of increasing protein diversity and is calculated to occur, at least at the transcriptional if not at the protein level, in >90% of human genes. Among the ARCs, human CD157 is the first family member to be described as being regulated by AS. What is intriguing about AS in CD157 is that it is not used to add to phenotypic diversity but, on the contrary, to maintain the status quo. By skipping exon 1b, the canonical proteoform of CD157 maintains the 1:1 sequence conservation and functional characteristics of its orthologues in mouse (NCBI accession number Q64277), rat (Q63072) and chicken (XP_420775). This seems to conform with observations that have emerged from large-scale proteomics experiments indicating that, despite AS being so widespread, most protein-coding genes have one dominant proteoform.

How did a novel BST1 exon arise during primate evolution? We can only speculate on the origins of BST1 exon 1b, which may have occurred following exonization of an intronic sequence, an event estimated to occur at the rate of 2.71 × 10⁻³/gene/million years. By comparative phylogenomics, exon 1b orthologous sequences were not detected in other mammals, apart from within the primate suborder Anthropoidea (monkeys and apes including human), suggesting that the birth of exon 1b may have occurred after the separation of the anthropoids from their sister group, the tarsiers, approximately 67 mya. Exon 1b origin is estimated to have occurred after the anthropoid/tarsier separation approx. 67 mya (orange dot).

Figure 5. Origin, distribution and dynamic evolution of BST1 exon 1b in primate phylogeny. Phylogenetic tree of primate species harboring exon 1b within the context of the BST1 gene. Species are: gibbon (Nomascus leucogenys, Nle); orang-utan (Pongo abelii, Pab); human (Homo sapiens, Hsa); chimpanzee (Pan troglodytes, Pta); bonobo (Pan paniscus, Ppa); gorilla (Gorilla gorilla, Ggo); snub-nosed monkey (Rhinopithecus roxellana); drill (Mandrillus leucomalicus); olive baboon (Papio anubis); rhesus macaque (Macaca mulatta); cynomolgus monkey (Macaca fascicularis); sooty mangabey (Cercopithecus atys); night monkey (Aotus nancymaeae); squirrel monkey (Saimiri boliviensis); common marmoset (Callithrix jacchus); prosimians (lemur, loris, tarsier).

Alignment of the exon 1b sequence is shown, highlighting the correlation between exon sequence evolution and phylogenetic lineage. Species with boxed names have genome sequences that are published or in progress. Chimpanzee (Pan troglodytes) and bonobo (Pan paniscus), part of the Pan genus lineage, both show AT mutation at the exon 1b 5' splice site (5' ss). An 11 bp sequence of exon 1b is deleted in the ape lineage after its separation from Old World monkeys approximately 28–25 mya. Exon 1b origin is estimated to have occurred after the anthropoid/tarsier separation approx. 67 mya (orange dot).
antibody, in addition to the unaltered, or indeed enhanced, ability to act as an adhesion protein and signalling molecule. However, the deleterious effect of exon 1b insertion on CD157-mediated NADase activity appears to be remarkable and probably explains the use of AS as a means of maintaining the catalytically active form of CD157. Examination of the 3D model presented here shows that the exon 1b-encoded 15 aa insert is predicted to form a structurally disordered, hydrophilic loop protruding from the membrane-distal surface of the protein. Downstream of the insertion, the CD157-002 polypeptide structure prediction appears to overlap with the 3D structure of CD157-001. On the one hand, this conservation of structure would not be expected to lead to changes in ECM component binding. On the other hand, the model does not offer an immediate explanation for the loss of the NADase activity in CD157-002. It is not obvious how the loop could hinder substrate entry or binding into the catalytic cleft, but further investigation into the capacity of CD157-002 to bind NAD$^+$ is warranted. It is also not clear what structural and functional effects the insert could have on the first N-glycosylation site, located three amino acids downstream of the insertion.

The use of AS to maintain a catalytically active form of CD157 raises the question why maintain a catalytically inactive variant? Among the possibilities, CD157-002 may create new protein-protein interactions, as suggested by recent findings that protein isoform pairs share less than 50% of their interactions. Alternatively, CD157-002 may be a more effective ECM receptor, a possibility not excluded by the experimental evidence presented here, without the NAD-consumption activity which may impact upon NAD$^+$ homeostasis. Last but not least, CD157-002 may possess additional functional characteristics that were undetectable with the assays performed here. Whatever the distinctive functional properties of CD157-002, the fact remains that, so far, the two CD157 proteoforms always seem to be co-expressed. Therefore, when performing functional experiments that evaluate the role of CD157 in physiological or pathological conditions, the outcome will be determined by the relative contribution of the two proteoforms.

In conclusion, BST1 is proving to be a prime example of a gene that has deployed all of the major mechanisms that increase phenotypic diversity, from its origin by tandem gene duplication and diversification, to multifunctionality to AS with inclusion/skipping of a novel cassette exon. Our findings add a new level of complexity to all the analyses in which CD157 is being evaluated for its relevance to human disease or its suitability as a therapeutic target in selected diseases. Perhaps sibling rivalry between CD157 and CD38 has just begun.

**Methods**

**Antibodies, cell lines and reagents.** Anti-CD157 mAb (SY/11B5, IgG1) was produced in-house and affinity purified on protein G. The sheep anti-human CD157 polyclonal antibody (AF4736, R&D Systems). Horseradish peroxidase (HRP)-conjugated anti-mouse IgG and anti-sheep IgG, and anti-β-actin-HRP were from Santa Cruz Biotechnology. Alexa Fluor-488-F(ab')$^2$-GoMlgG was from Invtrogen (Life Technologies). FITC-labelled F(ab')$_2$RoMlgG [F(ab')$_2$]-RoMlgG-FITC] was from Jackson ImmunoResearch.

Human cell lines THP1 and U937 (myelomonocytic leukaemias) and HeLa (cervical cancer) were originally obtained from the American Type Culture Collection. Cells were maintained at 37°C in RPMI 1640 with 2.5 μg/ml amphotericin B, 100-unit penicillin, 100 μg/ml streptomycin and 10% FCS. Cell cultures were routinely checked for the absence of mycoplasma contamination by PCR. Fibronectin, collagen I, vitronectin, crystal violet, propidium iodide, 1,N$^6$-ethenoadenine dinucleotide (ε-NAD$^+$) and nicotinamide guanine dinucleotide (NGD$^+$) were from Sigma Aldrich.

**Isolation of human PMN.** PMN were isolated to ≥95% purity from peripheral blood of healthy volunteers, as described. All participants gave written informed consent. All experiments were carried out in accordance with relevant guidelines and regulations, approved by the Institutional Review Board (Department of Medical Sciences, University of Torino).

**Analysis of gene expression by RT-PCR.** Total RNA was extracted from 2 × 10$^6$ PMN collected in TRIzol® reagent (Invitrogen) using Direct-Zol® RNA MiniPrep Kit (Zymo Research) according to the manufacturer’s instructions. Total RNA tissue samples were purchased from Invitrogen. Using oligo-dT primers, 2 μg RNA were reverse transcribed with M-MLV Reverse Transcriptase (ThermoFisher Scientific), according to the manufacturer’s instructions. cDNAs from selected cell types and tissues were from Human MTC™ Panel II (Clontech Laboratories). The cDNAs were amplified (30 cycles) using KAPA2G Fast HotStart DNA polymerase (Kapa Biosystems). Each cycle consisted of denaturation at 94°C for 10 s, annealing for 10 s and extension at 72°C for 1 s, followed by a final extension at 72°C for 40 s. For co-amplification of BST1-001 and -002 transcripts, primers were: Ex1/F 5′-ACACTTGGGGACATCTTCC-3′; Ex2/R 5′-GGGATAGAGTGCTGGACAGA-C3′. For BST1-002-specific amplification, primers were: Ex1b/F 5′-GCAGAGAACACCCTGTGTTA-3′; Ex2/R 5′GAACATCGCTAGGGCACATA-3′. Annealing temperature (Ta) was 57°C for both primer pairs. β-actin primers were: forward 5′-ATGGATGTGATGATATCGCCGCG-3′; reverse 5′-CTAGAAAGCTTGGGTCGATGACAG-3′. Ta was 55°C. Negative control reactions with only a single primer, with RNA without reverse transcription, or in the presence of RNase were performed. PCR products were analysed by agarose gel electrophoresis and visualized by staining with Midori Green (Nippon Genetics).

**PCR product sequencing.** PCR products were purified of unused primers and nucleotides by treatment with ExoSAP-IT® (Amersham-Pharmacia Biotech), followed by direct sequencing using the ABI PRISM® BigDye™ Terminator Cycle Sequencing kit (Applied Biosystems) and analysis on an ABI 3130 Genetic Analyzer with Sequencing Analysis software (Applied Biosystems).
Real-Time Quantitative PCR expression analysis. Real-Time quantitative RT-PCR (RT-qPCR) was carried out using an ABI_Prim 7500 Fast System (Applied Biosystems) with TaqMan Gene Expression Master Mix (Applied Biosystems) and Universal Probe Library (UPL) technology (Roche) according to the manufacturer’s instructions. Primers and UPL probes used for RT-qPCR were: for Bst1-001 transcript, forward 5’-CACACTGGCGGACATCTT-3′ and reverse 5’-TGTCAGGCACTCTATTCCC-3′, UPL no. 76; for Bst1-002 transcript, forward 5’-CAGCAGAGAACCACCTGTGTTA-3′; reverse 5’-TGTCAGGCACTCTATTCCC-3′, and UPL no. 76. Each sample was examined in triplicate, in three independent experiments. Experimental Ct values were normalized to the housekeeping gene GUSB, used as endogenous control (VIC-labelled predesigned TaqMan gene expression assays, GUSB, 4326320E, Applied Biosystems). The relative quantity of each transcript was calculated following normalization to GUSB using the conventional comparative (ΔΔCt) method.

Constructs and cell transfection. Plasmid constructs directing expression of canonical CD157 or CD157-002 were prepared by cloning the relative human full-length cDNAs into pCDNA3-puro, an in-house modified variant of the pCDNA3 eukaryotic expression vector (Invitrogen) that contains a puromycin resistance (puro') cassette. Plasmids were transfected into HeLa cells using Nanofectamin (PAA) according to the manufacturer’s protocol. HeLa cells transfected with empty pCDNA3 vector (indicated as mock) were used as controls. Stable transfectants were selected for 3 days in medium with 5μg/ml puromycin (SantaCruz Biotechnology) yielding ~100% transfected cells.

Flow cytometry and confocal microscopy. For flow cytometric analysis, cells (2 × 10^6) were incubated with SY/11B5 mAb (5μg/ml) for 30 min at 4°C followed by F(ab')2-GoMlgG-FITC for 30 min at 4°C. Fluorescence was analysed by means of a FACScanto flow cytometer using CellQuest software (BD Biosciences). Background mAb binding was estimated by means of isotype-matched control IgG. Three independent readings were obtained from separate experiments. For confocal microscopy, cells were grown on glass coverslips for 24 h, washed with HBSS and fixed with 2% paraformaldehyde (PFA). Coverslips were then blocked with 2% goat serum in HBSS for 30 min at RT, before incubation with anti-CD157 mAb and Alexa Fluor-488-F(ab')2-GoMlgG, and staining with propidium iodide. Samples were analysed using an Olympus FV300 laser scanning confocal microscope and Fluoview 300 software (Olympus Biosystems). Cells were imaged using a ×60 oil immersion objective (1.4 NA).

Cell adhesion assay. HeLa cell transfectants (3 × 10^4 in serum-free RPMI 1640) were plated in 96-well plates, previously coated with 10μg/ml fibronectin, collagen 1 or vitronectin and blocked with 0.2% BSA. After 10 min adhesion at 37°C, nonadherent cells were removed by washing. Adherent cells were fixed with 2% PFA and stained with crystal violet solution (0.5% wt/vol in H_2O) for 20 min, washed and solubilized in triton X-100 (1.0% solution in H_2O). The absorbance readings were recorded at 595 nm using a microplate reader (Infinite® 200 PRO). The absorbance of HeLa cell transfectants was compared with those of mock cells (basal HeLa cell adhesion) and expressed as fold increase of cell adhesion.

Wound healing assay. Cells were seeded in 24-well plates until confluence and a scratch was then made across the monolayer with a sterile pipette tip. The remaining monolayers were then gently rinsed with serum-free medium to remove detached cells and debris, and serum-free medium or medium with 10% FCS was added, as indicated. Wounds were photographed at the beginning of the experiment (t = 0) and at 24 h, and at least 20 randomly chosen distances across the wound for each time point were measured and the mean calculated. The distance travelled by the cells was determined as the difference between the width of the cell-free area after 24 h and at time 0. Wound images were acquired using an Olympus Biosystems Microscope IX70, equipped with a UC-30 camera and CellF software (Olympus Biosystems).

Western blotting and phosphorylation assay. Total lysates of HeLa cell transfectants, PMN and U937 cells were obtained in ice-cold RIPA buffer supplemented with protease inhibitor cocktail (Sigma Aldrich). Where indicated, cell lysates were treated with N-glycosidase (PNGase F) (Sigma Aldrich) as previously described. Protein extracts (30μg/lane) were separated by: for Hela cells, 12% SDS-PAGE under non-reducing conditions and transferred to PVDF membranes (Millipore); for PMN and U937, by 4-12% gradient SDS-PAGE under reducing conditions and transferred to nitrocellulose membranes. After blocking with 5% milk, membranes were probed with anti-CD157 mAb or anti-CD157 polyclonal antibody followed by the appropriate HRP-conjugated anti-IgG, or with anti-β-actin-HRP. Immunoreactive bands were visualized by Westar ECL substrate (Cyanagen). For phosphorylation assays, HeLa cell transfectants were serum-starved overnight, and then incubated for 30 min at 37°C in fresh medium without or with 10% FCS. Cells were then placed on ice, treated with 1 mM pervanadate and lysed with RIPA buffer supplemented with 1 mM Na_3VO_4, 5 mM NaF, 50μg/ml aprotinin and leupeptin. Lysates were run on 10% SDS-PAGE under reducing conditions and transferred to PVDF membranes. Membranes were blocked (2% BSA) and probed with anti-phospho-FAK (Tyr-397), anti-phospho-AKT (Ser-493) and anti-phospho-ERK1/2 (Tyr-204/Tyr-187) (Santa Cruz Biotechnology), then stripped and reprobed with anti-FAK, anti-AKT and anti-ERK2 antibodies (Santa Cruz) followed by incubation with the appropriate HRP-conjugated antibodies. Immunoreactive bands were visualized by ECL. Images were acquired with a ChemiDoc XR+ System, and densitometric analysis was performed using Image Lab v5.2.1 software (Bio-Rad). Quantification of protein band density was performed for each individual protein according to the formula:

\[
\text{Phospho} = \frac{\text{protein density of HeLa transfectants}}{\text{protein density of HeLa mock}}
\]

\[
\text{Total protein density of HeLa transfectants} = \frac{\text{total protein density of HeLa mock}}{\text{Total protein density of HeLa mock}}
\]
**NAD⁺ glycohydrolase and NGD⁺ cyclase enzymatic activities.** HeLa cell transfectants cells were detached in 2 mM EDTA, centrifuged 1 min at 4,000 rpm, suspended at 4 × 10⁵ in 10 mM potassium phosphate buffer, pH 7.4, and transferred to 96-well optical bottom black plates (NUNC) at 1 × 10⁵ cells/250 µl/well, in triplicate. Substrate was added to intact cells and the reaction was measured over 95 min at 37°C. Ecto-NAD glycohydrolase activity was measured continuously by fluorescence-based assay using 40 µM ε-NAD⁺ as substrate. Cyclase activity was evaluated by cyclization of the NAD⁺ surrogate substrate NGD⁺ to its fluorescent derivative cyclic GDP ribose (cGDPR) whose production was measured continuously in the presence of 100 µM NGD⁺ as substrate. Hydrolysis of ε-NAD⁺ and cyclization of NGD⁺ were determined by measuring the accumulation of fluorescent reaction products over time using an Infinite® F200 (Tecan) fluorescence plate reader with excitation and emission wavelengths of 300 nm (BW 20 nm) and 430 nm (BW 35 nm), respectively. The rate of ε-ADPR or cGDPR production was determined from the slope of fluorescence increase. Data are reported in arbitrary units (A.U.) of fluorescence intensity at 430 nm.

**Bioinformatics.** To identify homologous sequences, the human BST1 exon 1b nucleotide and protein sequences were used to query genomic databases by BLASTN and BLASTP searches, respectively. In most cases, retrieved sequences were annotated with information on their placement within a gene and chromosomal location. The exon 1b homologs included here were all located in orthologous positions within the BST1 gene of the species in question. Sequence alignments was obtained with MAFFT version7 (http://mafft.cbrc.jp/alignment/server/). Protein structure homology modelling was performed using the SWISS-MODEL program using PDB structure 1ISI of CD157 complexed with etheno-NAD⁺. The primate species tree was generated using the Common Tree tool at https://www.ncbi.nlm.nih.gov/taxonomy. Databases used for multispecies analysis of the BST1 gene and proteins were: (NCBI https://www.ncbi.nlm.nih.gov/), Ensembl (www.ensembl.org/) and UCSC (https://genome.ucsc.edu/) genome browsers. Intrinsic disorder in the human CD157 protein was predicted using IUPred (http://iupred.enzim.hu/)4. Putative branch points were predicted using Human Splicing Finder (http://www.umd.be/HSF/), version 2.4.1.30 and Sroogle (http://sroogle.tau.ac.il/)51.

**Statistical methods and data analysis.** Unless otherwise indicated, results are expressed as means ± sem. Comparisons between two groups were carried out using the Student’s t test. One-way analysis of variance (ANOVA) with Dunnett’s post hoc analysis was used for multiple comparisons. Statistical analyses were performed using GraphPad Prism 7 software (San Diego, CA). For all analyses, differences were considered significant at p < 0.05 (⁎ p < 0.05, ⁎⁎ p < 0.01, ⁎⁎⁎ p < 0.001, ⁎⁎⁎⁎ p < 0.0001 versus control; ns, not significant).

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

**References**

1. Ferrero, E., Lo Buono, N., Horenstein, A. L., Funaro, A. & Malavasi, F. The ADP-ribosyl cyclases–the current evolutionary state of the ARCs. *Front Biosci (Landmark Ed)* 19, 986–1002 (2014).

2. Ishihara, K. History and Perspectives in Bone Marrow Stromal Cell Antigen-1 (BST-1/CD157): a relative of ADP-ribosyl cyclase. *Biology of Blood Myeloid Cells* 1, 41–49 (2013).

3. Todd, R. F. 3rd, Roach, J. A. & Arnaout, M. A. The modulated expression of Mo5, a human myelomonocytic plasma membrane substrate. Hydrolysis of fluorescent reaction products over time using an Infinite® F200 (Tecan) fluorescence plate reader with excitation and emission wavelengths of 300 nm (BW 20 nm) and 430 nm (BW 35 nm), respectively. The rate of ε-ADPR or cGDPR production was determined from the slope of fluorescence increase, Data are reported in arbitrary units (A.U.) of fluorescence intensity at 430 nm.

4. Funaro, A. et al. ADP ribosyl cyclase activity of a novel bone marrow stromal cell surface molecule, BST-1. *FEBS Lett* 516, 184–188, https://doi.org/10.1016/S0014-5793(98)00563-8 (1999).

5. Funaro, A. et al. The BST1–integrin partnership controls transendothelial migration and adhesion of human monocytes. *Biol Chem* 388, 2985–2993, https://doi.org/10.1515/bchm.2017.248 (2017).

6. Lo Buono, N. et al. The CD157–integrin partnership controls transendothelial migration and adhesion of human monocytes. *Biol Chem* 388, 2985–2993, https://doi.org/10.1515/bchm.2017.248 (2017).

7. The primate species tree was generated using the Common Tree tool at https://www.ncbi.nlm.nih.gov/taxonomy. Databases used for multispecies analysis of the BST1 gene and proteins were: (NCBI https://www.ncbi.nlm.nih.gov/), Ensembl (www.ensembl.org/) and UCSC (https://genome.ucsc.edu/) genome browsers. Intrinsic disorder in the human CD157 protein was predicted using IUPred (http://iupred.enzim.hu/)4. Putative branch points were predicted using Human Splicing Finder (http://www.umd.be/HSF/), version 2.4.1.30 and Sroogle (http://sroogle.tau.ac.il/)51.

8. Ortolan, E. et al. Binding of CD157 protein to fibronectin regulates cell adhesion and spreading. *Front Biosci (Landmark Ed)* 14, 929–943 (2009).

9. Morone, S. et al. Overexpression of CD157 contributes to epithelial ovarian cancer progression by promoting mesenchymal differentiation. *PLoS One* 7, e43649, https://doi.org/10.1371/journal.pone.0043649 (2012).

10. Buckley, S. A. & Walter, R. B. Antigen-specific immunotherapies for acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program* 2015, 584–595, https://doi.org/10.1182/asheducation-2015.1.584 (2015).

11. Mirkowska, P. et al. Leukemia surfaceome analysis reveals new disease-associated features. *Blood* 121, e149–159, https://doi.org/10.1182/blood-2012-11-468702 (2013).

12. Yamanoto-Katayama, S. et al. Crystallographic studies on human BST-1/CD157 with ADP-ribose cyclase and NAD glycohydrolase activities. *J Mol Biol* 316, 711–723, https://doi.org/10.1016/j.jmb.2001.5386 (2002).

13. Buckley, S. A. & Walter, R. B. Antigen-specific immunotherapies for acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program* 2015, 584–595, https://doi.org/10.1182/asheducation-2015.1.584 (2015).

14. Morone, S. et al. Overexpression of CD157 contributes to epithelial ovarian cancer progression by promoting mesenchymal differentiation. *PLoS One* 7, e43649, https://doi.org/10.1371/journal.pone.0043649 (2012).

15. Ortolan, E. et al. CD157 enhances malignant pleural mesothelioma aggressiveness and predicts poor clinical outcome. *Oncotarget* 5, 6191–6205, https://doi.org/10.18632/oncotarget.2188 (2014).

16. Krujawa, C. et al. Targeting CD157 in AML using a novel, Fc-engineered antibody construct. *Oncotarget*. https://doi.org/10.18632/oncotarget.16060 (2017).

17. Buckley, S. A. & Walter, R. B. Antigen-specific immunotherapies for acute myeloid leukemia. *Hematology Am Soc Hematol Educ Program* 2015, 584–595, https://doi.org/10.1182/asheducation-2015.1.584 (2015).

18. Mirkowska, P. et al. Leukemia surfaceome analysis reveals new disease-associated features. *Blood* 121, e149–159, https://doi.org/10.1182/blood-2012-11-468702 (2013).

19. Yamanoto-Katayama, S. et al. Crystallographic studies on human BST-1/CD157 with ADP-ribose cyclase and NAD glycohydrolase activities. *J Mol Biol* 316, 711–723, https://doi.org/10.1016/j.jmb.2001.5386 (2002).

20. Hirata, Y. et al. ADP ribose cyclase activity of a novel bone marrow stromal cell surface molecule, BST-1. *FEBS Lett* 356, 244–248 (1994).
Acknowledgements

This work is dedicated to the memory of our coauthor Simona Morone. We thank Prof. Fabio Malavasi (University of Torino, Italy) for critically reading the manuscript. We also thank our colleagues Drs. Erika Ortolan and Elisa Giorgio (University of Torino, Italy) for their valuable help. This work was supported by grants from AIRC - Italian Association for Cancer Research (AIRC, IGI2014 I0.15968 to AF); from MIUR - the Italian Ministry for University and Scientific Research (Ex60% Projects to AF and to EF); and by Ricerca Sanitaria Finalizzata (GR – 2011-02348356 to NLB). SA is supported by a fellowship by Fondazione Giovanni Goria (Torino, Italy).

Author Contributions

N.L.B., S.M. and R.P. designed and carried out molecular biology and biochemistry experiments, and analysed data; C.M. performed qPCR assays; A.B. supervised the qPCR assays; A.G. and S.A. helped setup the transfectant cell model and performed cell biology experiments; A.R.V., S.G.R., M.Z. and J.S. performed enzymatic activity data; C.M. performed qPCR assays; A.B. supervised the qPCR assays; A.G. and S.A. helped setup the transfectant cell model and performed cell biology experiments; A.R.V., S.G.R., M.Z. and J.S. performed enzymatic activity

21. Hasan, M. A. et al. Scant Extracellular NAD Cleaving Activity of Human Neutrophils is Down-Regulated by sMLP via FPRL1. Korean J Physiol Pharmacol 18, 497–502, https://doi.org/10.4196/kjp.2014.18.6.497 (2014).
22. Fliegert, R., Gasser, A. & Guse, A. H. Regulation of calcium signalling by adenosine-based second messengers. Biochem Soc Trans 35, 109–114, https://doi.org/10.1042/BST0351019 (2007).
23. Podesta, M. et al. Concentrative uptake of cyclic ADP-ribose generated by BST-1+ stroma stimulates proliferation of human hematopoietic progenitors. J Biol Chem 280, 5343–5349, https://doi.org/10.1074/jbc.M408055200 (2005).
24. Yilmaz, O. H. et al. mTORC1 in the Paneth cell niche couples intestinal stem-cell function to calorie intake. Nature 486, 490–495, https://doi.org/10.1038/nature11163 (2012).
25. Higashida, H. et al. An immunohistochemical, enzymatic, and behavioral study of CD157/BST-1 as a neuroregulator. BMC Neurosci 13, 35, https://doi.org/10.1186/s12868-017-0350-7 (2017).
26. Lopatina, O. L., Furuhara, K., Ishihara, K., Salmina, A. B. & Higashida, H. Communication Impairment in Ultrasonic Vocal Repertoire during the Suckling Period of CD157 Knockout Mice: Transient Improvement by Oxytocin. Front Neurosci 11, 266, https://doi.org/10.3389/fnins.2017.00266 (2017).
27. Wang, S. et al. Association between bone marrow stromal cell antigen 1 gene polymorphisms and the susceptibility to Parkinson's disease: a meta-analysis. Neurosci Lett 599, 120–126, https://doi.org/10.1016/j.neulet.2015.05.026 (2015).
28. Satake, W. et al. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson’s disease. Nat Genet 41, 1303–1307, https://doi.org/10.1038/ng.485 (2009).
29. Lopatina, O. et al. Anxiety- and depression-like behavior in mice lacking the CD157/BST1 gene, a risk factor for Parkinson's disease. Front Behav Neurosci 8, 133, https://doi.org/10.3389/fnbeh.2014.00133 (2014).
30. Muraoka, O., Tanaka, H., Itoh, M., Ishihara, K. & Hirano, T. Genomic structure of human BST-1. Immunol Lett 54, 1–4 (1996).
31. Dong, C., Willerford, D., Alt, F. W. & Cooper, M. D. Genomic organization and chromosomal localization of the mouse Bp3 gene, a member of the CD38/ADP-ribosyl cyclase family. Immunogenetics 45, 35–43 (1996).
32. Ferrero, E., Saccucci, F. & Malavasi, F. The human CD38 gene: polymorphism, CpG island, and linkage to the CD157 (BST-1) gene. Immunogenetics 49, 597–604 (1999).
33. Koren, E., Lev-Maor, G. & Ast, G. The emergence of alternative 3′ and 5′ splice site exons from constitutive exons. PLoS Comput Biol 3, e95, https://doi.org/10.1371/journal.pcbi.0030095 (2007).
34. Dostie, Y., Zizun, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21, 3433–3434, https://doi.org/10.1093/bioinformatics/bti451 (2005).
35. Biastoch, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using experimental information. Nucleic Acids Res 42, W252–258, https://doi.org/10.1093/nar/gku340 (2014).
36. Zola, H., Swart B., Nicholson I. & Voss, E. Lysosome and Stromal Cell Molecules: The CD Markers. 289 (Wiley, 2007).
37. Lee, H. & Cyclic ADP-ribose and nicotinic acid adenine dinucleotide phosphate (NAADP) as messengers for calcium mobilization. J Biol Chem 287, 31633–31640, https://doi.org/10.1074/jbc.R112.349464 (2012).
38. Yamamoto-Katayama, S. et al. Site-directed removal of N-glycosylation sites in BST-1/CD157: effects on molecular and functional heterogeneity. Biochem J 357, 385–392 (2001).
39. Lund, F. E., Moutin, M. J., Muller-Steffiner, H. & Schuber, F. ADP-ribosyl cyclase and GDP-ribosyl cyclase activities are not always equivalent: impact on the study of the physiological role of cyclic ADP-ribose. Anal Biochem 346, 336–338, https://doi.org/10.1016/j.abb.2005.08.016 (2005).
40. Matlin, A. J., Clark, F. & Smith, C. W. Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6, 386–398, https://doi.org/10.1038/nrm1645 (2005).
41. Koren, H., Lev-Maor, G. & Ast, G. Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet 11, 345–355, https://doi.org/10.1038/nrg2776 (2010).
42. Abascal, F. et al. Alternatively Spliced Homologous Exons Have Ancient Origins and Are Highly Expressed at the Protein Level. PLoS Comput Biol 11, e1004325, https://doi.org/10.1371/journal.pcbi.1004325 (2015).
43. Ezkurdia, I. et al. Most highly expressed protein-coding genes have a single dominant isoform. J Proteome Res 14, 1880–1887, https://doi.org/10.1021/pr5012863 (2015).
44. Lev-Maor, G. et al. The “alternative” choice of constitutive exons throughout evolution. PLoS Genet 3, e203, https://doi.org/10.1371/journal.pgen.0030203 (2007).
45. Wang, W. et al. Origin and evolution of new exons in rodents. Genome Res 15, 1258–1264, https://doi.org/10.1101/gr.3929705 (2005).
46. Zhang, X. H. & Chasin, L. A. Comparison of multiple vertebrate genomes reveals the birth and evolution of human exons. Proc Natl Acad Sci USA 103, 13427–13432, https://doi.org/10.1073/pnas.0603042103 (2006).
47. Pagani, F. & Baralle, F. E. Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet 5, 389–396, https://doi.org/10.1038/nrg1327 (2004).
48. Yang, X. et al. Widespread Expansion of Protein Interaction Capabilities by Alternative Splicing. Cell 164, 805–817, https://doi.org/10.1016/j.cell.2016.01.029 (2016).
49. Chini, C. C., Tarrago, M. G. & Chini, E. N. NAD and the aging process: Role in life, death and everything in between. Mol Cell Endocrinol. https://doi.org/10.1016/j.mce.2016.11.003 (2016).
50. Desmet, F. O. et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37, e67, https://doi.org/10.1093/nar/gkp215 (2009).
51. Schwartz, S., Hall, E. & Ast, G. SROOGLE: webserver for integrative, user-friendly visualization of splicing signals. Nucleic Acids Res 37, W191–192, https://doi.org/10.1093/nar/gkp320 (2009).
52. Rogers, J. & Gibbs, R. A. Comparative primate genomics: emerging patterns of genome content and dynamics. Nature reviews. Genetics 15, 347–359, https://doi.org/10.1038/nrg3707 (2014).
assays and analysed data; A.F.P. was involved in data interpretation; E.F. conceived the study, provided ideas and supervision, performed phylogenetic analysis and prepared the figures; E.F. and A.F. analysed and interpreted all the data and wrote the manuscript. All authors approved the manuscript.

**Additional Information**

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-16184-w.

**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.

© The Author(s) 2017