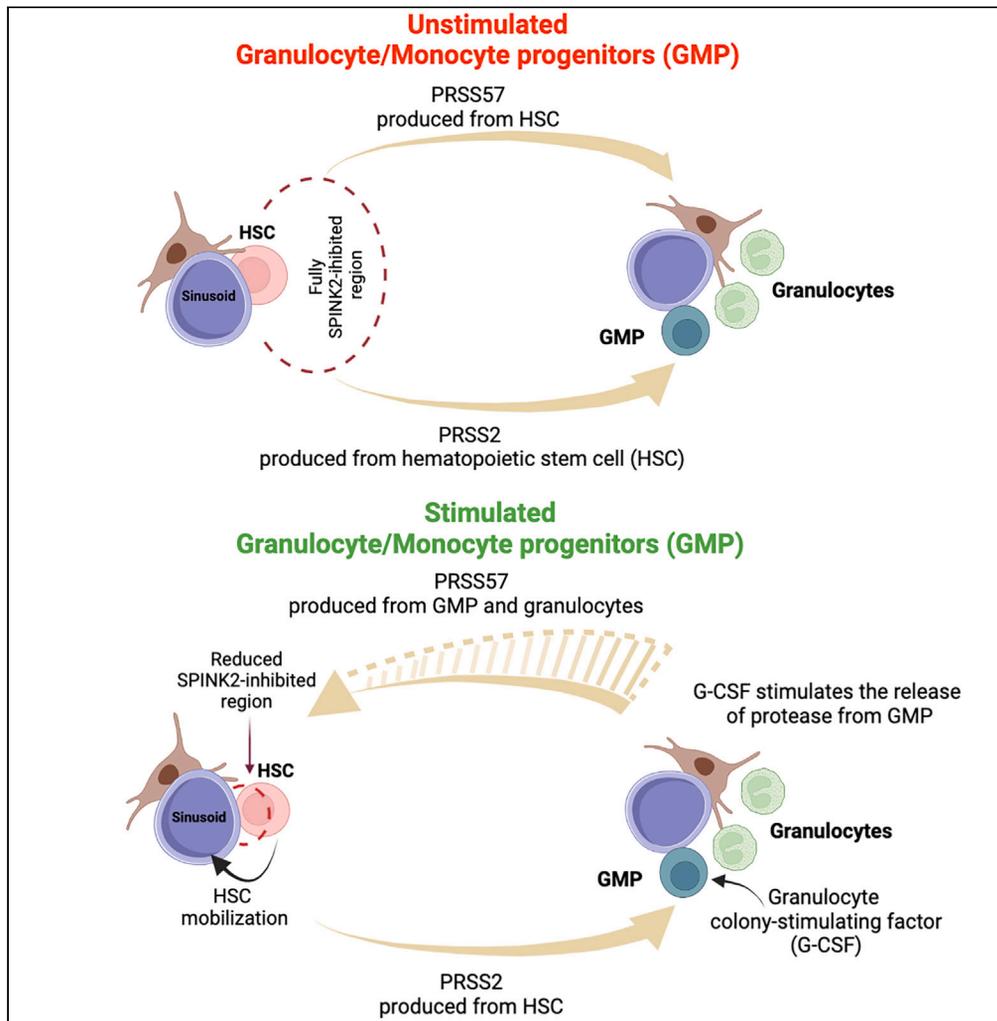# iScience

Article

Temporary serine protease inhibition and the role of SPINK2 in human bone marrow

Vincenza Barresi,
Virginia Di Bella,
Luca Lo Nigro,
Anna Provvidenza
Privitera, Paola
Bonaccorso,
Chiara Scuderi,
Daniele Filippo
Condorelli

vincenza.barresi@unict.it (V.B.)
daniele.condorelli@unict.it
(D.F.C.)

Highlights

Physiological SPINK2
expression in adult bone
marrow and in mobilized
CD34+ cells

SPINK2 expression in
haematological cancers

Kinetic properties of
trypsin inhibition by
SPINK2

Physiological meaning of
the temporary enzyme
inhibition properties of
SPINK2

Article

Temporary serine protease inhibition and the role of SPINK2 in human bone marrow

Vincenza Barresi,[1,5,*] Virginia Di Bella,[1,5] Luca Lo Nigro,[2,3] Anna Provvidenza Privitera,[1] Paola Bonaccorso,[2,3] Chiara Scuderi,[1] and Daniele Filippo Condorelli[1,4,6,*]

SUMMARY

Protease temporary inhibitors are true substrates that bind the catalytic site with high affinity but are slowly degraded, thus acting as inhibitor for a defined time window. Serine peptidase inhibitor Kazal type (SPINK) family is endowed with such functional property whose physiological meaning is poorly explored. High expression of SPINK2 in some hematopoietic malignancies prompted us to investigate its role in adult human bone marrow. We report here the physiological expression of SPINK2 in hematopoietic stem and progenitor cells (HSPCs) and mobilized cluster differentiation 34 (CD34)[+] cells. We determined the SPINK2 degradation constant and derived a mathematical relationship predicting the zone of inhibited target protease activity surrounding the SPINK2-secreting HSPCs. Analysis of putative target proteases for SPINK2 revealed the expression of PRSS2 and PRSS57 in HSPCs. Our combined results suggest that SPINK2 and its target serine proteases might play a role in the intercellular communication within the hematopoietic stem cell niche.

INTRODUCTION

The serine peptidase inhibitor Kazal type (SPINK) protein family is formed by serine protease inhibitors characterized by the Kazal domain. This domain is composed of a relatively conserved sequence of 50–60 amino acid residues, including 6 typical cysteine residues forming three pairs of disulfide bonds that stabilize its conformation. The Kazal domain includes the so-called "active site loop" able to bind the catalytic site of the serine protease and to inhibit it. In humans the SPINK family is formed by 10 members. The first member of the family (SPINK1) was discovered by Kazal et al. (1948)[1] in bovine pancreatic extracts and is known as "pancreatic secretory trypsin inhibitor". SPINK1 is produced by acinar cells in the exocrine pancreas and released into the pancreatic ducts. It has been suggested that its main role is to inhibit prematurely activated trypsin both in the acinar cell and during its journey in the pancreatic juice.[2–7] The second member of the SPINK family (SPINK2) has been discovered as an inhibitor of acrosin, a serine protease localized in the acrosome of spermatozoa and involved in the so-called acrosomal reaction. Analysis at the tissue or organ level indicates that SPINK2 expression is restricted to testis, epididymis, and seminal vesicles and the inhibition of prematurely activated acrosin is the proposed main role of SPINK2 in such tissues.[8–10] This functional role is in agreement with the observation that loss-of-function mutations of SPINK2 are associated with azoo-oligospermia and male infertility.[9,10]

Our group has recently reported that SPINK2 is overexpressed in a subset of pediatric acute myeloid leukemias (AMLs) and such overexpression is associated with a form of primary chemoresistance to the standard treatment for this form of leukemia.[11] Indeed, upregulation of SPINK2 has been previously observed in adult AML patients, and its expression has been reported as an indicator of poor prognosis.[12] Moreover, it has been reported that SPINK2 is significantly elevated in several leukemia cell lines.[13]

With this in mind we considered it is crucial to answer some basic questions about the expression and function of SPINK2 in order to understand its pathophysiological role in AML. The first question is whether SPINK2 expression in AML derives from an aberrant heterotopic gene activation in malignant leukemia cells or from the physiological expression of this protein in bone marrow hematopoietic cells that are precursors to leukemia cells. Analysis of global gene expression databases shows that SPINK2 expression is negligible in bone marrow samples (www.proteinatlas.org/ENSG00000128040-SPINK2/tissue)[14] suggesting that bone marrow expression, if present, should be limited to a specific minor cell subpopulation.

[1]Department of Biomedical and Biotechnological Sciences, Section of Medical Biochemistry, University of Catania, 95123 Catania, Italy

[2]Cytogenetic-Cytofluorimetric-Molecular Biology Lab, 95123 Catania, Italy

[3]Center of Pediatric Hematology-Oncology, Azienda Policlinico – San Marco, 95123 Catania, Italy

[4]Senior author

[5]These authors contributed equally

[6]Lead contact

*Correspondence: vincenza.barresi@unict.it (V.B.), daniele.condorelli@unict.it (D.F.C.)

https://doi.org/10.1016/j.isci.2023.106949

In the present study we analyze single-cell RNA sequencing (scRNA-seq) databases providing clues on the presence and identity of such subpopulation and validate the physiological presence of SPINK2 in bone marrow hematopoietic cells by expression studies in cluster differentiation 34 (CD34) + purified cells. The second question is related to the quantitative features of the type of enzymatic inhibition exerted by SPINK2 on serine proteases. Our analysis of kinetic features of such inhibition in an *in vitro* system and in mathematical models suggests the hypothetical physiological role of SPINK2 in bone marrow and sheds light on the possible biological consequence of its retained expression in leukemia cells. The third question is the identity of the bone marrow serine protease(s) that are physiological targets of SPINK2. On the basis of gene expression data obtained in CD34$^+$ cells, we provide, for the first time, evidence for the presence of some putative targets, including an isoform of trypsin, in hematopoietic stem cells (HSCs).

RESULTS

High levels of SPINK2 in CD34$^+$ bone marrow cells

Preliminary analysis of scRNA-seq data from human blood and bone marrow cell types[15–17] revealed an un-detectable level of SPINK2 mRNA in several bone marrow cell types such as endothelial cells, fibroblasts, adipocytes, osteoblasts, macrophages, and mature myeloid and lymphoid cells[18] (Figures S1–S3, related to Figures 1 and 5). A low level of SPINK2 was present in a small percentage (<5%) of osteoclasts. On the contrary, a relatively high expression was observed in hematopoietic stem and progenitor cells (HSPCs). Therefore, a detailed analysis of the transcription level for SPINK2 in HSPCs, defined as CD34$^+$, within the bone marrow was performed using the scRNA-seq data by Setty et al. (2019)[19] available through the "Human Cell Atlas Data Portal" (https://data.humancellatlas.org/explore/projects/091cf39b-01bc-42e5-9437-f419a66c8a45). The following CD34$^+$ HSPCs of the human bone marrow were identified by RNA sequencing (RNA-seq) data in agreement with the standard model of hematopoiesis[19]: HSC, hematopoietic multipotent progenitor (HMP), common myeloid progenitor (CMP), granulocyte-monocyte progenitor (GMP), dendritic progenitor (DP), megakaryocyte progenitor (MP), erythroid progenitor (EP), and common lymphoid progenitor (CLP) (Figure 1A). As shown in Figures 1B and 1C, SPINK2 mRNA levels (expressed as count per million, CPM) and percentages of SPINK2- expressing cells are very high in HSCs, HMPs, and CMPs (550–800 CPM and 90–100%, respectively). Expression of other members of SPINK family was absent or negligible except for SPINK9. SPINK2 expression was much higher than SPINK9 in HSCs, HMPs, CMP, and GMP, while SPINK9 expression was higher than SPINK2 in MPs, EPs, and CLPs (Figure 1C).

In order to verify the presence of SPINK family members in bone marrow HSPCs of a common laboratory mouse model, we analyzed scRNA-seq data by Nestorowa et al. (2016),[20] made available through the plat-form "Single Cell Expression Atlas" (www.ebi.ac.uk/gxa/sc/home). Spink2 and Spink12 (the putative mouse ortholog of human SPINK9[21]) were not detectable in HSPCs of mouse C57BL/6. Among other members of the SPINK family, only low level of Spink10 (no known human ortholog) and very low levels of Spink4 and Spink14 were observed as shown in Figure 1D.

Enrichment of SPINK2 transcript and protein in human CD34$^+$ cells

CD34 is a transmembrane glycoprotein expressed in HSPCs. Granulocyte colony-stimulating factor (G-CSF) was used to stimulate the mobilization of CD34$^+$ cells from the human bone marrow to the peripheral blood, and mobilized CD34$^+$ cells were positively selected with paramagnetic beads coated with anti-CD34 antibodies from blood mononuclear cells collected by leukapheresis and enriched by Ficoll-Hypaque density gradient centrifugation. By quantitative real-time PCR (qRT-PCR) we evaluated the transcript levels of SPINK2 in the purified CD34$^+$ cell population and in the remaining blood population of mononuclear cells from which CD34$^+$ cells have been separated (indicated as CD34$^-$cells). As shown in Figure 2A, we confirmed a significant enrichment of SPINK2 transcripts in purified CD34$^+$ cells compared to CD34$^-$cells (600-fold enrichment). ELANE (neutrophil elastase) transcript was used as negative control for CD34$^+$ enrichment since it is highly expressed in both CD34$^-$ differentiated neutrophil granulocytes and in CD34$^+$ GMPs.

Expression of SPINK2 protein was confirmed in human CD34$^+$ HSPC by publicly available proteomics data[22] and by western blot in CD34$^+$ purified cells. As shown in Figures 2B and 2C, abundant and specific expression of SPINK2 can be observed in CD34$^+$ HSPCs in agreement with the high level of SPINK2 transcript (Figure 2A). A strong homology among SPINK1 and SPINK2 proteins is showed in Figures 2D and 2E. 63 out of 84 amino acids are perfectly conserved (red), particularly highlighted in Kazal domain including
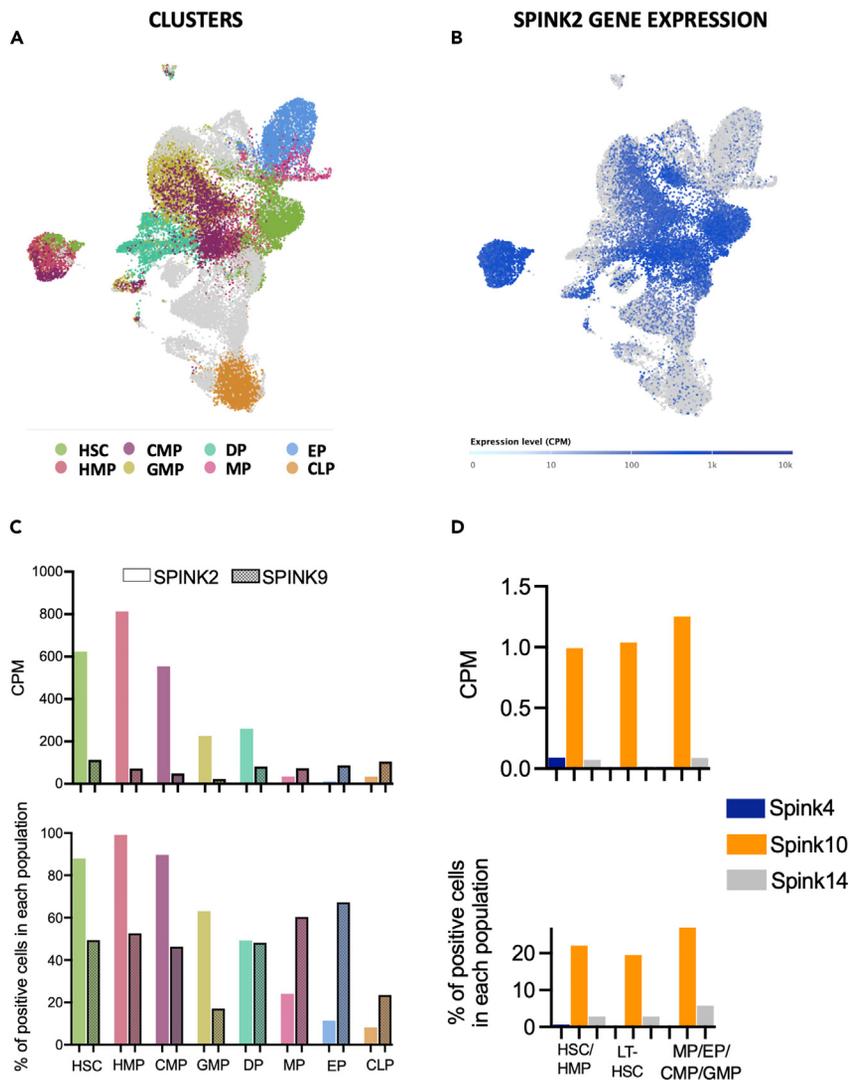
Figure 1. SPINK2 is highly expressed in CD34+ bone marrow cells

(A) Uniform manifold approximation and projection (UMAP) graph shows the CD34+ cell clusters within human bone marrow as annotated by Setty et al.[19]

(B) UMAP graph showing SPINK2 expression in CD34+ cells.

(C) SPINK2 and SPINK9 expression values are reported as averaged normalized CPM (top graph), and percentage of SPINK2 or SPINK9 positive cells in different population of CD34+ cells within the bone marrow (bottom graph). A cell is considered positive if normalized CPM value is > 0. Numbers of analyzed cells (n) in each population are the following: HSC, hematopoietic stem cell (n = 4690); HMP, hematopoietic multipotent progenitor (n = 4306); CMP, common myeloid progenitor (n = 2328); GMP, granulocyte-monocyte progenitor (n = 3713); DP, dendritic progenitor (n = 2075); MP, megakaryocyte progenitor (n = 507); EP, erythroid progenitor (n = 3463); CLP, common lymphoid progenitor (n = 3237);

(D) Averaged normalized CPM (left graph) and percentage of SPINK positive cells (right graph) in different populations of hematopoietic cells within the bone marrow of mouse C57BL/6. LT-HSC, Long Term-HSC (n = 216); HSCs/HMPs, hematopoietic stem and progenitor cells (n = 852); MP/EP/CMP/GMP, megakaryocyte progenitor/erythroid progenitor/common myeloid progenitor/granulocyte-monocyte progenitor (n = 851).

the six cysteines that stabilize the inhibitor (cysteines are located at position 32, 39, 47, 58, 61, and 79 in SPINK1 and at position 36, 44, 52, 63, 66, and 84 in SPINK2).

SPINK2 expression in human tissues and cancer cell lines

Using data deposited in Cell Model Passport (https://cellmodelpassports.sanger.ac.uk/), we analyzed the expression level of SPINK2 in 1,293 cancer cell lines, grouped for 26 tissue origins. As shown in Figure 3A,

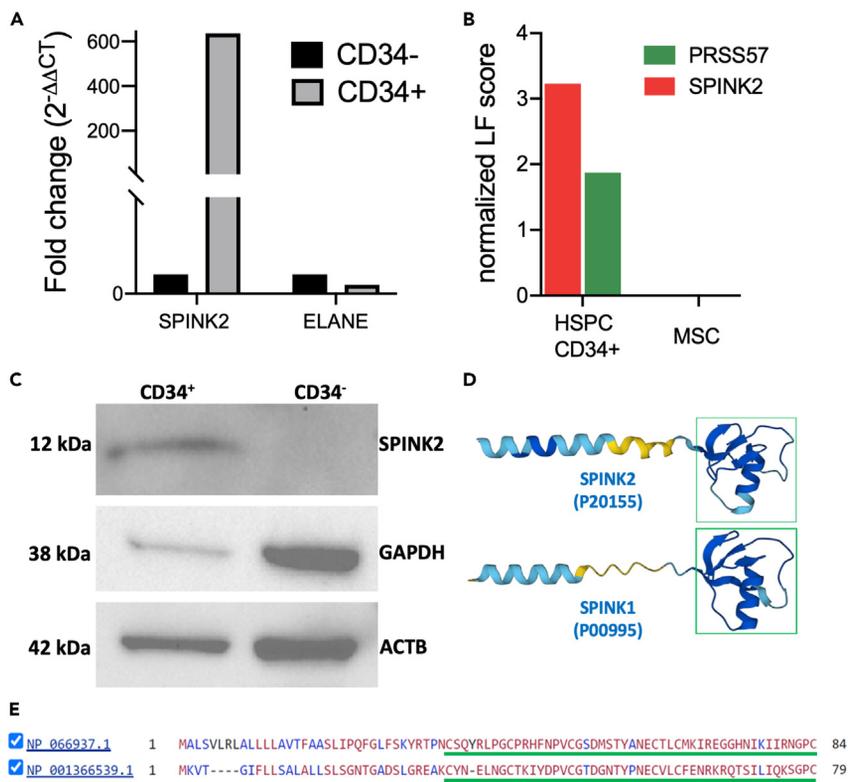**Figure 2. SPINK2 transcript and protein are highly expressed in CD34$^+$ bone marrow cells**

(A) qRT-PCR by $2^{-\Delta\Delta CT}$ method confirms high levels of SPINK2 in purified CD34$^+$ cells. ELANE, typically expressed in differentiated neutrophil granulocytes and granulocyte-monocyte progenitors, has been evaluated as negative control.

(B) SPINK2 and PRSS57 are highly expressed at protein level in CD34$^+$ HSPCs. Protein intensity of SPINK2 and PRSS57 is expressed as normalized label-free (LF) score in CD34$^+$ HSPC (Hennrich et al.[22]). Negative control results in bone marrow mesenchymal stromal cells (MSC) are also shown.

(C) Western blot showing SPINK2 enrichment in CD34$^+$ cells. GAPDH and ACTB were the loading control.

(D) SPINK2 (P20155) and SPINK1 (P00995) structures have been downloaded from AlphaFold Protein Structure Database developed by DeepMind and EMBL-EBI (https://alphafold.ebi.ac.uk/entry/P20155; https://alphafold.ebi.ac.uk/entry/P00995).[23,24] Kazal domain is highlighted in green square.

(E) SPINK2 (NP_066937, 84 amino acids) and SPINK1 (NP_001366539) protein alignment was performed by Constrain-based Multiple Alignment tool (Cobalt, https://www.ncbi.nlm.nih.gov/tools/cobalt/cobalt.cgi), and as showed in Figure 2E the method highlights in red color highly conserved aminoacidic positions based on the relative entropy threshold of the residue, while blue color indicates lower conservation. Kazal domain is underlined (green) and is located among amino acids in position 36 and 84 and includes 6 typical cysteines forming three pairs of disulfide bonds that stabilize its conformation.

the higher levels of SPINK2 expression were found in hematopoietic-lymphoid and testis tissues. Focusing on hematopoietic-lymphoid cancer cells, Figures 3B–3D show SPINK2 expression for each cell line derived from AML (n = 42), acute lymphoblastic leukemia (ALL, n = 45), and lymphoma (n = 86). A subpopulation of cell lines shows a high level of SPINK2 expression in those three hematological cancer types. Indeed, 16.3% (14/86), 21.4% (9/42), and 33.3% (15/45) of cancer cell lines show averaged transcripts per million (TPM) > 10 for lymphomas, AML, and ALL, respectively (Figures 3B–3D).

By qRT-PCR we analyzed SPINK2 expression in three AML cell lines (HL-60, K562, and Kasumi-1) and one ALL cell line (Jurkat). HCT116, a model of colorectal cancer, was used as control and calibrator. As shown in Figure 3E, SPINK2 was highly expressed only in the Jurkat cell line, as expected from RNA-seq data reported above.

Trypsin inhibition by recombinant human SPINK2

As a model system for the study of the inhibitory properties of SPINK2, we used a porcine trypsin modified by reductive methylation in order to resist auto-proteolytic digestion. Recombinant human SPINK2 (Met1-Cys84)
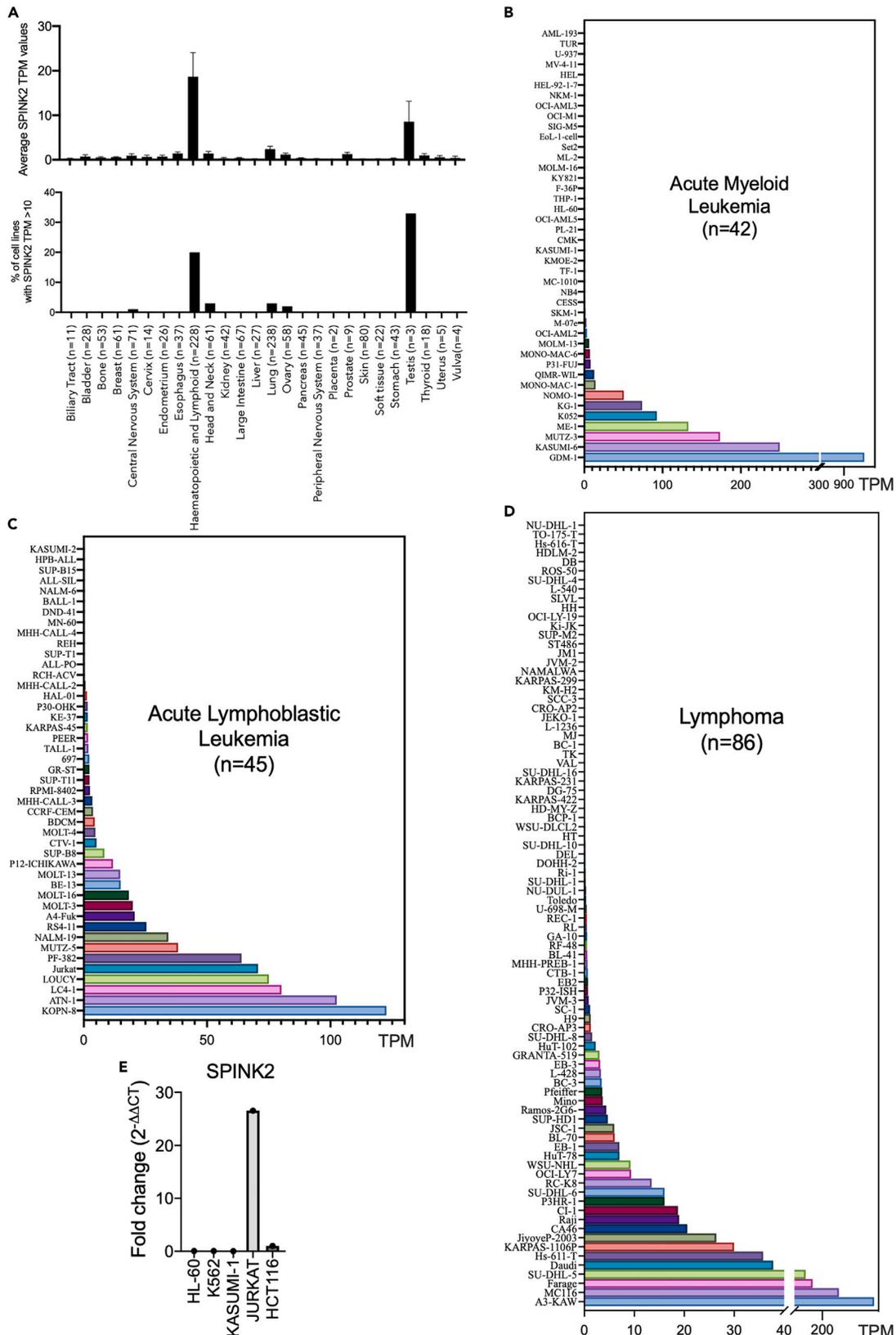
Figure 3. SPINK2 expression in human tissues and cancer cell lines

(A) SPINK2 expression in 1,293 cancer cell lines assembled for tissue of origin in 26 groups. Averages of TPM ± SD and percentage of positive cell lines (>10 TPM) are shown in top and bottom histograms, respectively.

(B) SPINK2 expression levels in AML cell lines (n = 42); (C) in ALL cell lines (n = 45); and (D) in Lymphoma cell lines (n = 86).

(E) qRT-PCR confirmed SPINK2 expression in Jurkat cell line, a model of ALL, while other leukemia cell lines (HL-60, K562, KASUMI-1) did not express SPINK2. HCT116, a colon cancer cell line, was used as calibrator for fold change calculation.

was expressed with the Fc region of human immunoglobulin G1 (IgG1) at the C terminus in HEK293 cells, and the mature protein (N-terminal at Gln24 after cleavage of signal peptide) was purified.

Determinations of initial reaction rates were performed by incubating trypsin in absence and in presence of different concentrations of recombinant SPINK2. Preliminary titration experiments with different enzyme concentrations were performed in order to obtain a proper 2X dilution series of SPINK2. Since SPINKs belong to the class of tight-binding inhibitors,[7,13] characterized by an apparent affinity near the concentration of enzyme in the biochemical assay, the inhibitory constant (Ki) was determined using Morrison's quadratic equation for tight-binding inhibition.[25] Utilizing previously determined K_m values of porcine trypsin (2.45 mM) for the N-Benzoyl-Phe-Val-Arg-p-nitroanilide hydrochloride (FVR-NA) substrate,[13] SPINK2 exhibited Ki values in the nanomolar range (Ki 11.2 nM; 95% confidence interval: 7.9–16.1 nM) (Figures 4 A and B).

Members of the SPINK family are also considered temporary inhibitors.[7,26] In particular, temporary inhibition occurs when the inhibitor binds to the catalytic site as a competitive inhibitor but it is also a true substrate of its target protease, although stable enough to exert its inhibitory function before being degraded.

The more general and simpler scheme is the following:

$$Trypsin + SPINK2 \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} Trypsin^* SPINK2 \overset{^{SPINK2}k_{cat}}{\rightarrow} Trypsin + degraded\ SPINK2$$

In the first reaction enzyme and inhibitor bind and dissociate based on k_{on} or k_{off}, respectively. In the second reaction the rate of degradation of the inhibitor depends on the specific rate constant of trypsin for SPINK2 degradation ($^{SPINK2}k_{cat}$). Szabo et al.[7] have recently shown for SPINK1 that k_{on} and k_{off} show highly divergent values (3 x 10^6 M^{-1} s^{-1} and 54.8 x 10^{-6} sec^{-1}, respectively) and the resulting Kd is in the subnanomolar range, thus confirming the behavior as tight-binding inhibitors of members of the SPINK family. In order to obtain an estimate of $^{SPINK2}k_{cat}$ we collected data of prolonged progress curves relating the concentration of the product formed (dependent variable) at each time (independent variable). Figure 4C shows the accumulation of product during a 17 h-period in the absence and in the presence of SPINK2. Results obtained with two different SPINK2/Trypsin molar ratios are reported (Figure 4C).

Experimental data are then compared to those calculated by a mathematical model based on a combination of steady-state and time-course enzyme kinetic parameters.[27] We incorporated a SPINK2 degradation rate by trypsin in such model ($^{SPINK2}k_{cat}$), thus allowing a direct comparison of the "product versus time" curves in the presence of a significant degradation rate ($^{SPINK2}k_{cat} > 0$; temporary inhibition) or in its absence ($^{SPINK2}k_{cat} = 0$; persistent inhibition). Here we denominate this model as the temporary inhibitor model or TI-Model (see details in STAR Methods). In the graph of Figure 4D, we report the calculated data obtained with TI-Model, showing the different curves obtained modeling temporary or persistent inhibition. By constraining some parameters (fixed total trypsin concentration of 3.2 nM, experimentally determined $^{FVR-NA}V_{max}$ and $^{FVR-NA}K_m$ of trypsin, experimentally determined apparent K_m of trypsin in the presence of SPINK2, restricted range of 1–15 nM for $^{SPINK2}Ki$ value) and by fitting to the TI-model, it was possible to estimate the $^{SPINK2}V_{max}$ (2.3 x 10^{-7} mmol/min) and the $^{SPINK2}k_{cat}$ (1.2 x 10^{-3} sec^{-1}). Calculated results in Figure 4D have been obtained with the fitted $^{SPINK2}k_{cat}$. A 61000-fold ratio $^{FVR-NA}k_{cat}/^{SPINK2}k_{cat}$ can be so calculated.

As shown in Figure 4E, in the TI-Model the half-life of SPINK2 increases by increasing the initial concentration of the temporary inhibitor. This behavior is a consequence of the introduction in the TI-Model of the Michaelis and Menten relationship for SPINK2 degradation by trypsin: at increasing SPINK2/Ki ratios the reaction progressively approximates to a pseudo-zero order kinetics, and the half-lives depend on initial concentrations of SPINK2 (Figure 4E).
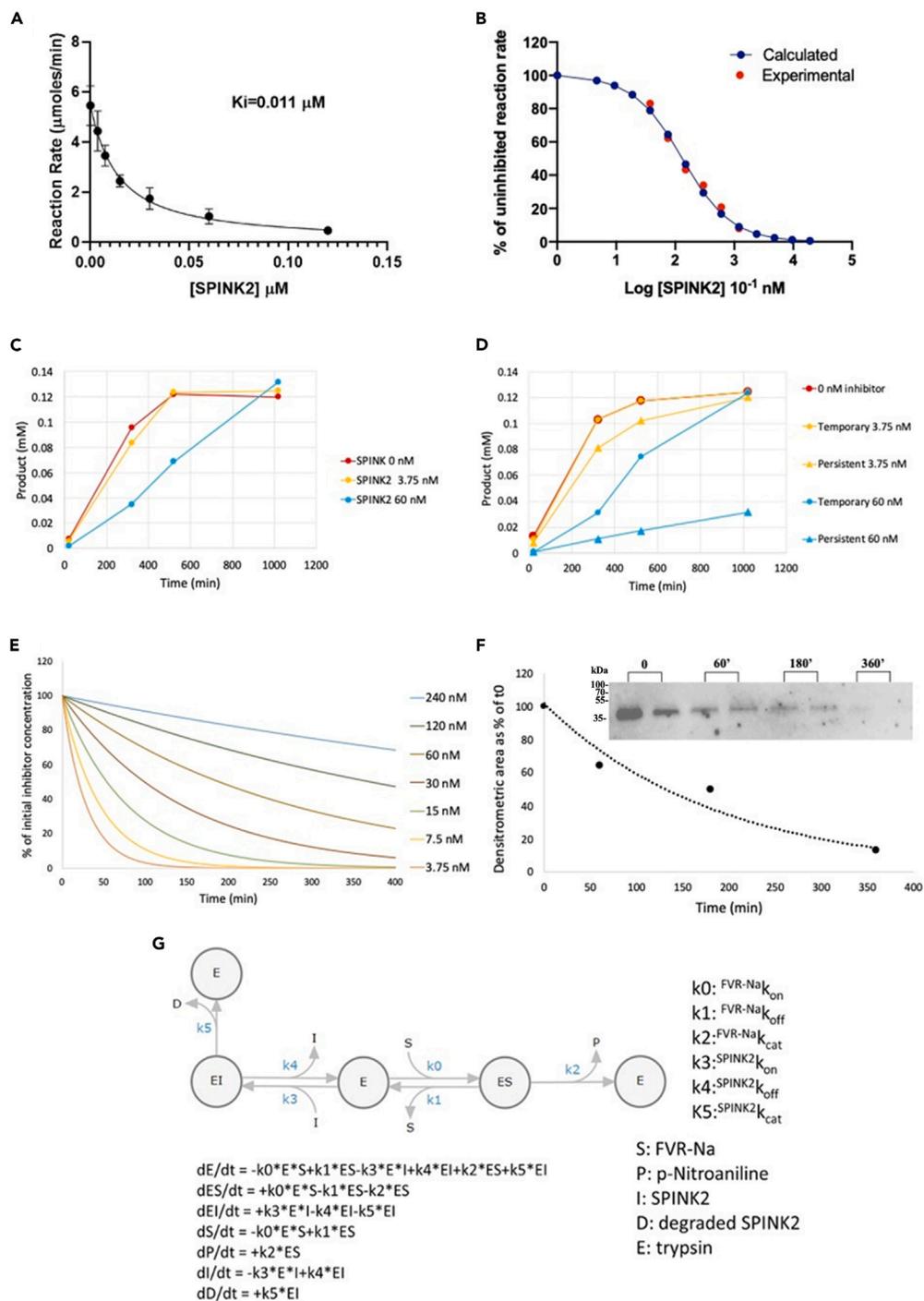
Figure 4. Kinetic properties of trypsin inhibition by SPINK2

(A) Graph showing experimental initial reaction rates at different concentrations of SPINK2 (range 3.75–120 nM) and fixed concentration of total enzyme (0.0032 μM) and initial substrate (125 μM); K_m was constrained at 2,450 μM. Ki was determined by fitting experimental data to Morrison's equation by GraphPad Prism 8.0.2. Results are expressed as averages $\pm$ SEM of two independent experiments.

(B) Initial reaction rates expressed as % of uninhibited reaction rate versus Log_{10} concentration of SPINK2. Blue lines and points are calculated values with a Ki = 0.011 μM; red points are experimental values.

(C) Experimental "product vs. time" curves; enzymatic reactions were followed for 17 h by monitoring spectrophotometrically the conversion of the substrate FVR-NA in *p*-nitroaniline at 410 nm. Trypsin (3.2 nM) was

Figure 4. *Continued*

incubated with two different SPINK2 concentrations (3.75 and 60 nM) corresponding to an inhibitor/enzyme ratio of 1.2 and 19, respectively.

(D) Calculated data obtained with TI-Model, showing the different "product vs. time" curves obtained with temporary and persistent inhibition; (E) Graph showing the "% of initial SPINK2 concentration" vs. Time obtained by TI-Model with K_i = 3 nM and increasing concentration of SPINK2. A $^{SPINK2}k_{cat}$ of 1.2×10^{-3} sec^{-1} and a fixed concentration of 3.2 nM trypsin were used in the simulation.

(F) *In vitro* degradation of recombinant SPINK2 by trypsin at different times. Each incubation was performed in duplicate. Averages of densitometric areas of western blotting bands (shown in the inset) are expressed as percentages of value at time zero (t0).

(G) Kinetic reaction scheme for enzyme activity, competitive enzyme inhibition, and inhibitor degradation. Differential rate equations and abbreviations for kinetic constants (k0-k5) and chemical species are also reported in the figure.

In order to confirm the estimates of SPINK2 degradation rates by trypsin, we performed direct experiments by incubating trypsin and SPINK2 and measuring the amount of remaining SPINK2 at different incubation times by western blot analysis (Figure 4F). A half-time of 125 ± 5 min can be calculated for 60 nM concentration of SPINK2 comparable with the 180 min predicted by the TI-Model.

Modeling of trypsin-inhibitory action of SPINK2 by differential equations and fitting of experimental progress curve

We applied the web tool ENZO[28] to generate seven differential equations for SPINK2-inhibited trypsin according to the kinetic reaction scheme reported in Figure 4G and to fit the coefficients of differential equations to the progress curves of product concentration.

We used fixed initial concentrations of the various chemical species (as better specified in the STAR Methods section), fixed values of three kinetic constants (k0, k1, k2) related to non-inhibited trypsin activity as predetermined in initial velocity enzymatic assays, and two kinetic constants (k3, k4) related to SPINK2 association to and dissociation from trypsin, using values previously reported by Szabo et al.[7] for SPINK1 ($^{SPINK1}k_{on} = 3 \times 10^6$ mol^{-1} s^{-1}; $^{SPINK1}k_{off} = 5.48 \times 10^{-5}$ sec^{-1}). Only the k5 constant, corresponding to $^{SPINK2}k_{cat}$, was left unfixed, and different initial values, ranging from 1.6×10^{-4} to 8.3×10^{-1} sec^{-1}, were tested for fitting. All the iteration procedures converged with a fitted k5 value of 1.12×10^{-3} sec^{-1}, which is in good agreement with the value calculated in the previous section (1.2×10^{-3} sec^{-1}). Although limitations in the determination of the reaction rate constants by modeling with differential equations and fitting of progress curves[29,30] should be taken into account, the agreement of the $^{SPINK2}k_{cat}$ value with the experimental data on SPINK2 degradation by trypsin, reported in the previous subsection, provides a further support to such estimate.

Putative SPINK2-inhibited serine proteases in CD34+ bone marrow cells

In order to identify candidate targets of SPINK2, serine proteases expressed by HSPCs were analyzed using the scRNA-seq data by Setty et al.[19] available through the "Human Cell Atlas Data Portal" (https://data.humancellatlas.org/explore/projects/091cf39b-01bc-42e5-9437-f419a66c8a45). Firstly, we analyzed the expression of human serine protease genes reported in MEROPS database (www.ebi.ac.uk/merops/).[31] After removal of pseudogenes, we selected 224 human genes belonging to the following clan/family (the number of proteins is reported in brackets): PA/S1 (149), SB/S8 (10), SB/S53 (1), SC/S9 (48), SC/S10 (3), SC/S28 (3), SC/S33 (24), SE/S12 (2), SF/S26 (5), SH/S21 (1), SJ/S16 (2), SK/S14 (1), SK/S41 (3), SO/S74 (1), SP/S59 (2), SR/S60 (8), and ST/S54 (9). Validated normalized CPMs expressed in CD34+ bone marrow cells, as well as uniform manifold approximation and projection (UMAP) and T-SNE graphs, have been obtained for 178 genes (Table S1). In the latter group we selected 71 genes encoding for secreted protein according to the definition of the Human Protein Atlas (www.proteinatlas.org) further revised by a literature search, as indicated in Table S1.

To further narrow down the list of potential target proteases, we focused on enzymes expressed in HSCs at a level higher than 5% of SPINK2 value. After excluding the TF gene, encoding serotransferrin, because this protein does not bear protease activity and was only included in MEROPS database due to sequence homology, the following serine proteases (PRSSs) were selected: PRSS2, PRSS57, PRSS27, DPP7, and PLAT (for detailed data see Table S2). As shown in Figure 5A, PRSS2, the anionic form of trypsin, was highly expressed in HSCs, HMP, and CMP. For comparison, the expression values of PRSS1, the cationic form of trypsin, and PRSS3 (also known as mesotrypsin) are also reported in Figure 5A. PRSS1 is absent in CD34+ bone marrow cells, while a low but detectable expression of PRSS3 was observed in HSCs, HMP,

and CMP. On the contrary, PRSS27, PLAT, and DPP7 are more homogenously expressed among different types of HSPCs. (Figure 5B). PRSS27 encodes for marapsin, a tryptic serine protease acting on synthetic tetrapeptide substrates but with no general protease activity and with an unusual resistance to inhibitors such as aprotinin, serum serpins, and soybean trypsin inhibitor.[32] It has been suggested that the active site is partially blocked and does not allow access to large substrates and inhibitors. DPP7 (Dipeptidyl Peptidase 7) gene encodes for a serine aminopeptidase that cleaves dipeptides from the N-terminal of proteins that have a proline at the penultimate position.[33] The substrate specificity of DPP7 makes unlikely its role as target for SPINK2. PLAT (plasminogen activator, tissue type) gene encodes for a secreted serine protease, called tPA, whose main role is the conversion of the proenzyme plasminogen to plasmin, a fibrinolytic enzyme. Indeed, it has been reported that tPA, through upregulation of MMP9, is able to control the degradation of important molecule of the hematopoietic stem cell niche (HSCN), such as the membrane-bound kit ligand (stem cell factor or SCF) and the stromal cell-derived factor 1 (SDF1 or CXCL12) expressed by bone marrow stromal stem cells.[34,35] However, tPA is inhibited with high affinity and irreversibly by Plasminogen Activator Inhibitor Type 1 (PAI-1), a secreted serpin encoded by SERPINE1 gene,[36] thus making unlikely a role for SPINKs in its regulation.

PRSS57 encodes for one of the serine proteases (neutrophil serine protease 4 [NSP4]) produced and released from neutrophil granulocytes, together with neutrophil elastase (ELANE), cathepsin G (CTSG), and proteinase 3 (PRTN3). The expression values of all neutrophil serine proteases in different subtypes of CD34$^+$ HSPCs are reported in Figure 5C. As expected, all four neutrophil proteases are highly expressed in CMPs and GMPs. However, differently from the other neutrophil proteases, PRSS57 is also highly expressed in the most undifferentiated stages (HSC and HMP): 68% of HSCs and 90% of HMPs express PRSS57. Moreover, a relevant expression of PRSS57 is also observed in more restricted progenitors MPs and EPs (Figure 5C).

Considering previous results, we validated the expression of PRSS2 and PRSS57 in purified mobilized CD34$^+$ cells. As shown in Figure 5D, we detected a significant enrichment of both transcripts in CD34$^+$ cells compared to CD34$^-$ cells (147-fold and 9-fold for PRSS2 and PRSS57, respectively). PRSS1 transcript levels are reported as negative control.

In conclusion, due to their substrate specificity and inhibitor sensitivity, PRSS2 and PRSS57 appear to be the more probable candidate serine proteases expressed in the hematopoietic niches. The SPINK2/PRSS57 or SPINK2/PRSS2 ratios in different subpopulations of HSPCs are shown in Figure 5E. A SPINK2/PRSS2 expression ratio of 2–5.4 was observed in early undifferentiated hematopoietic cells (HSC, HMP) while the percentage of PRSS2-expressing cells was negligible (<0.5%) in more differentiated progenitors (GMP, MP, EP). A SPINK2/PRSS57 expression ratio of 2.2–2.3 was observed in early undifferentiated hematopoietic cells (HSC, HMP) and 0.3–0.6 in more differentiated progenitors (GMP, MP, EP).

In Figure 5F the percentages of SPINK2+ cells coexpressing PRSS2 or PRSS57 are reported for each CD34$^+$ subpopulation. Combined with results in Figure 5A, these data show that PRSS2+ cells represent only a small subpopulation of SPINK2-expressing CD34$^+$ HSCs, while PRSS57 is expressed in the large majority of the latter cell type. However, the majority of PRSS2+ or PRSS57+ HSCs and HMCs coexpress SPINK2 (Figure 5G). Moreover, the majority of PRSS2+ cells coexpress PRSS57 (HSCs: 62%, HMP: 81%; CMP: 96%; GMP: 92%; DP: 47%; MP: 100%; EP: 100%; CLP: no PRSS2+ CLP).

Expression of PRSS57 protein was confirmed by proteomics profiling of human bone marrow cells.[22] As shown in Figure 2B abundant expression of PRSS57 protein can be observed in CD34$^+$ HSPCs, supporting the results obtained at transcript level (Figure 5C).

In Figure 6 we compared the expression of SPINK2, PRSS2, and PRSS57 in CD34$^+$ cell types of human bone marrow. The expression of the other neutrophil serine proteases (ELANE, CTSG, PRTN3) is also reported for comparison. Expression value of a gene in a specific CD34$^+$ cell type is expressed as percentage of its maximal value among all different CD34$^+$ cell subtypes. SPINK2 shows its maximal value in HMPs, PRSS2 in HSCs, and PRSS57 in GMPs.

The temporary inhibition-diffusion (TI-D) theory for SPINK2 action in HSCN

The definition of the kinetics properties of the temporary inhibition exerted by SPINK2 on trypsin, the demonstration of its expression in HSCs, and the discovery that candidate SPINK2-inhibited serine
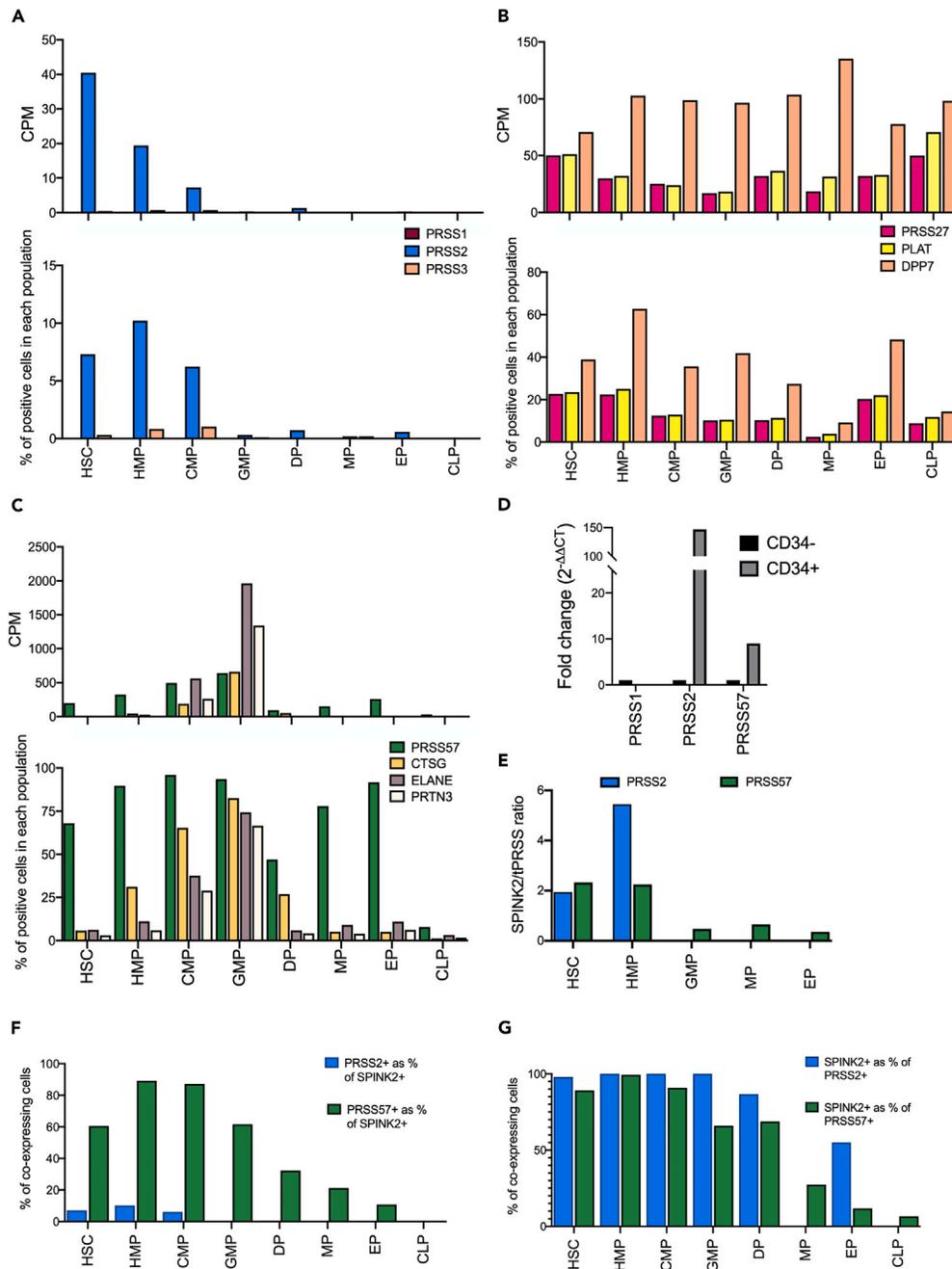
Figure 5. Expression of putative tPRSS in human bone marrow

Averages of normalized CPM values (top graph) and percentage of positive cells (bottom graph) in different CD34[+] cell subpopulations for (A) PRSS1, PRSS2, and PRSS3; (B) PRSS27, PLAT, and DPP7; and (C) PRSS57, CTSG, ELANE, and PRTN3.

(D) qRT-PCR analysis of PRSS1, PRSS2, and PRSS57 transcripts in purified mobilized CD34[+] blood cells. CD34[−]cells were used as calibrator for fold change calculation.

(E) SPINK2/PRSS57 or SPINK2/PRSS2 ratios in different subpopulations of HSPCs.

(F) Percentage of PRSS2+ and PRSS57+ cells in SPINK2+ cell population.

(G) Percentage of SPINK2+ cells in PRSS2+ cell population and in PRSS57+ cell population.
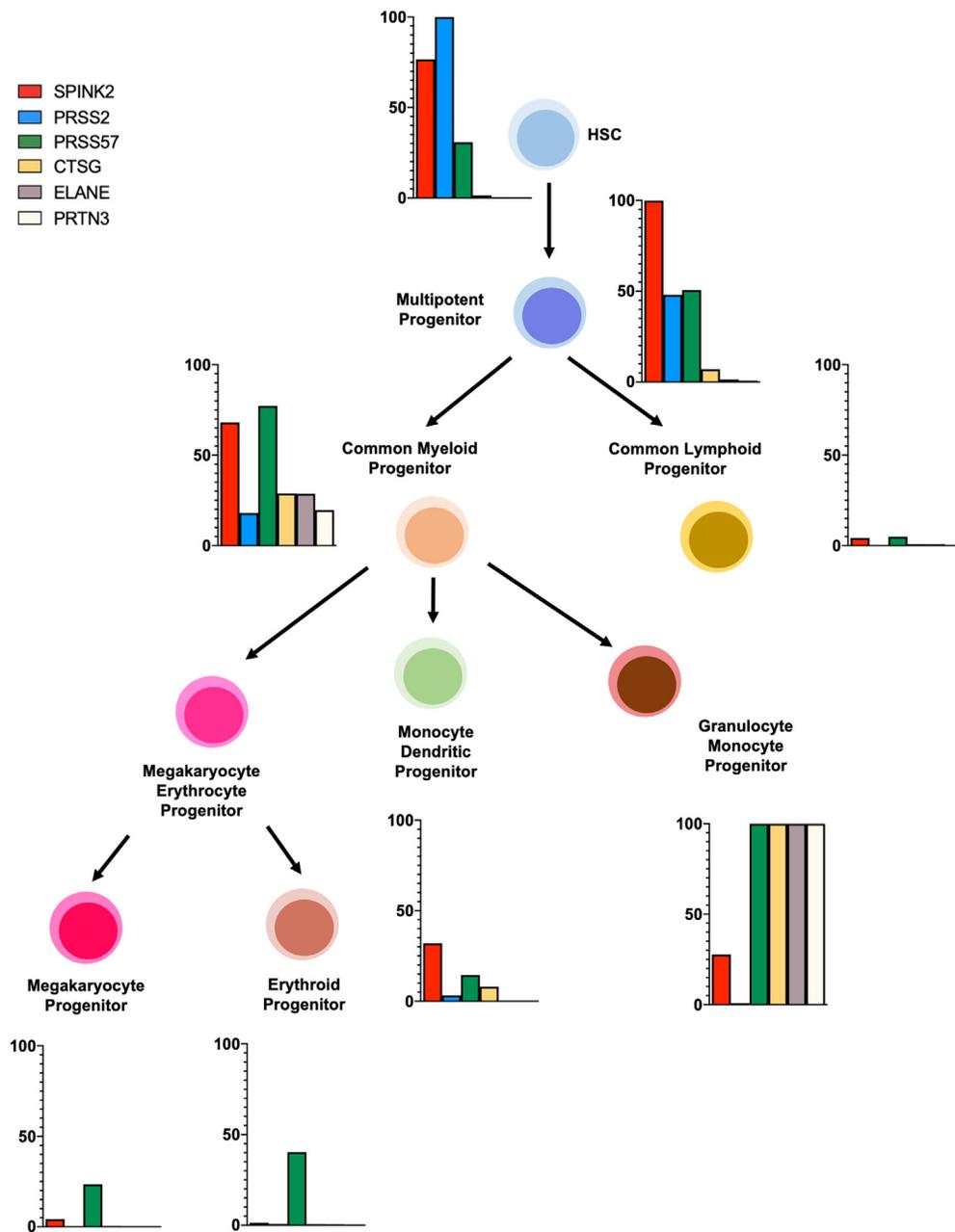
Figure 6. Differentiation tree of hematopoietic stem and progenitor cells

A graph reporting the transcript level of each gene expressed as percentage of their maximal value among CD34[+] cell types (normalized CPM value of gene A in cell type Z/max TPM value of gene A among all cell types x 100) is shown close to each cell type.

proteases, such as PRSS2 and PRSS57, are also expressed in HSCs prompted us to delineate a hypothesis, called TI-D theory, on the physiological role of this protein in bone marrow. Indeed, a well-accepted idea is that members of the SPINK family avoid a premature activation of target serine proteases both intracellularly and extracellularly. However, the temporary inhibitory properties and the production of both target protease and inhibitor by the same cell type suggested an additional function for SPINK2 beyond the simple protective role against inappropriate early activation of target proteases. This functional hypothesis assumes that the activation of the zymogen to active enzyme and the formation of enzyme/inhibitor complex are rapid and complete in a thin pericellular layer surrounding the producing HSCs. For the purposes of the model, the exact site of enzyme activation and complex formation is not relevant, i.e., intracellular or
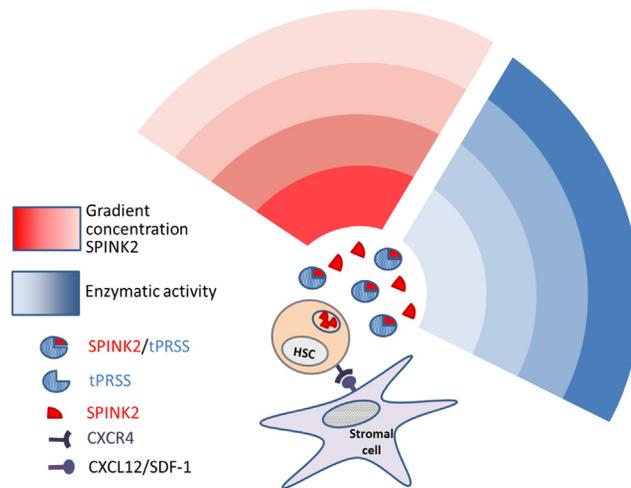
Figure 7. Schematic drawing illustrating the TI-D theory for SPINK2 role in hematopoietic stem cell niche

extracellular after secretion, but the diffusion of the secreted proteins far from the producing cell. Therefore, we hypothesize that diffusion generates extracellular steady-state gradients[37] of SPINK2 and its complexes with target protease. The gradient originates from the HSCs located in specialized microenvironments, called HSCN, scattered in the bone marrow.[38] The core of a single HSCN is assumed as composed of a single HSC or a cluster of few HSCs attached to leptin receptor-positive mesenchymal stromal cells (Figure 7). The generation of steady-state gradients by diffusing SPINK2 and its putative target serine protease (tPRSS) was modeled by the synthesis-diffusion-degradation (SDD) model,[39,40] as fully described in STAR Methods. Briefly, the specific model is based on the following assumptions: (i) a localized source, identified with HSCs in an HSCN, producing SPINK2 and tPRSS with fluxes $^{SPINK2}J_0$ and $^{tPRSS}J_0$, respectively; (ii) SPINK2 and tPRSS diffusion with a diffusion coefficient D; and (iii) a uniform general linear protein degradation rate k' equal for both SPINK2 and tPRSS. The aim of the model was to highlight the functional difference between a simulated persistent SPINK2 (not degraded by its tPRSS) and the real temporary one (degraded by its tPRSS). The specific SPINK2 degradation by tPRSS was modeled by introducing an additional degradation constant called k''. k'' was set equal to zero to simulate a persistent SPINK2, whereas values of k'' for temporary SPINK2 were calculated on the basis of concentration of SPINK2 at the source boundary (x = 0) using the mathematical relationship determined for the model porcine trypsin (Figure 4E). The last assumption of the model is that $k'' \gg k'$ within the biologically significant range of 200 μm from the HSC.

The SDD model provided the values of [tPRSS] and [SPINK2] at different distance x from the source, and the corresponding concentrations of the complex enzyme-inhibitor, [EI], were calculated by the Morrison quadratic equation, as described in STAR Methods. The results obtained using diffusion coefficient D ranging from 1 to 10 $μ^2$/sec[41,42] and a ratio [I]/[E] = 2 at the source boundary are reported in Figure 8A. By using the trypsin kinetic parameters ($^{SPINK2}k_{cat}$, $^{S}k_{cat}$, K_m), determined *in vitro* and reported in the previous section, it was possible to obtain an estimate of enzyme activity per unit volume at different distances from the source cell (Figure 8A). A region of fully inhibited tPRSS activity surrounds the source cells, and its size increases with increasing values of D. The estimated size of such region (40–110 μm) is compatible with the hypothesis that the delayed disinhibition of secreted tPRSS, due to temporary inhibition by SPINK2, allows the physiological protease action at a fixed distance from the producing HSC. The size of the region of fully inhibited tPRSS activity was largely dependent on the ratio [I]/[E] (i.e., [SPINK2]/[tPRSS]) at the source boundary, as shown in Figure 8B.

In previous simulations the steady-state concentration gradients for SPINK2 and tPRSS were simulated by the same degradation constant and the same SDD equation. The basic idea behind this choice is that SPINK2 and tPRSS were co-secreted by the same source cells and underwent a similar generic degradation process. We tested also two additional conditions: 1) a fixed homogeneous concentration of tPRSS along the entire diffusion path and 2) an inverse gradient of tPRSS by modifying the SDD Equation 5 (see STAR Methods). In both conditions we kept the ratio [I]/[E] constant (=2) at the SPINK2 boundary source. As
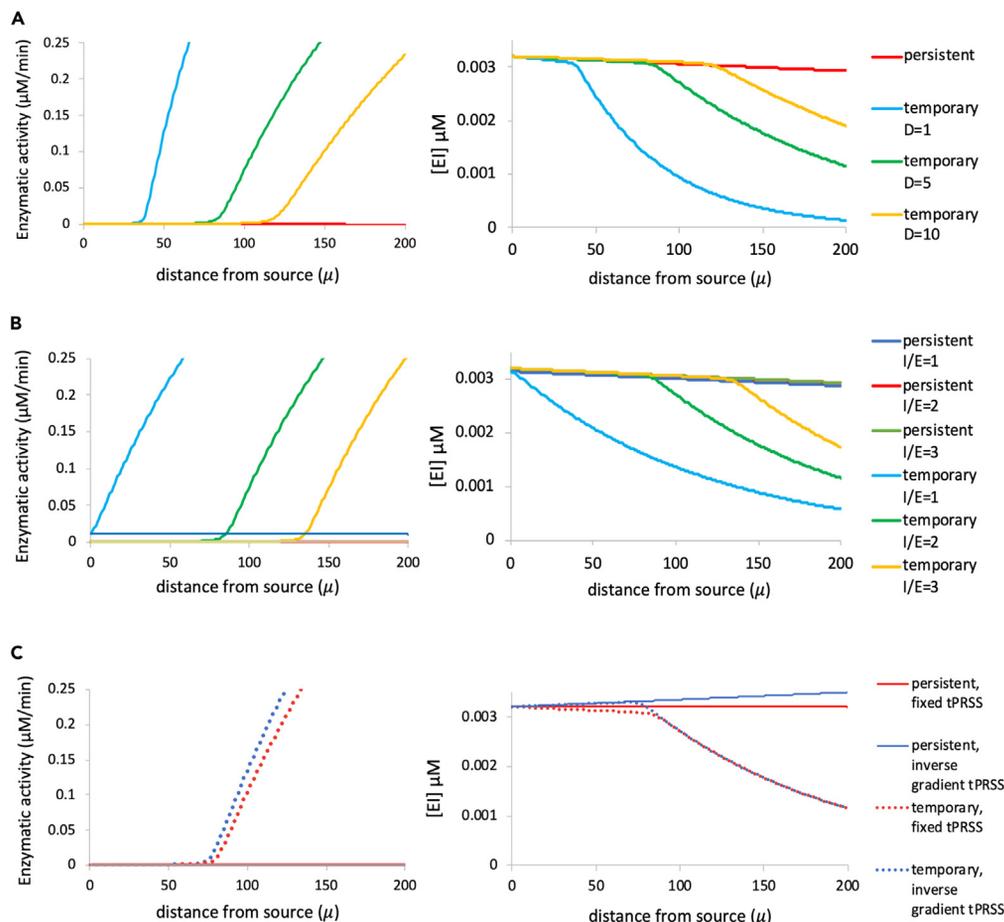
Figure 8. The temporary inhibition-diffusion model for SPINK2 in hematopoietic stem cell niche

(A) The concentration of the complex tPRSS SPINK2, indicated as [EI] and shown in top graphs, and the velocity of conversion of protein substrate (^{S}v), shown in bottom graphs, were calculated at different distances from the source cell using different values of D (μ^2/sec), as indicated above each plot. Results obtained for temporary or persistent inhibitor are shown in each plot. A SPINK2/tPRSS ratio = 2 is used in these simulations.

(B) The velocity of conversion of protein substrate (^{S}v) per unit volume was calculated at different distances from the source cell using different ratios [I]/[E] at the source boundary, as indicated above each plot. Results obtained for a temporary inhibitor or a persistent one are compared in each plot. A D value = 5 μ^2/sec is used in these simulations.

(C) The concentration of the complex [EI] is shown in top panels, and the velocity of conversion of protein substrate (Sv) per unit volume in bottom panels. A fixed total concentration of [E] is used in left panels, and an inverse concentration gradient in the right panels. Results obtained for a temporary inhibitor or a persistent one are compared in each plot. A SPINK2/tPRSS ratio = 2 at the boundary of inhibitor source and a D value = 5 μ^2/sec are used in these simulations.

shown in Figure 8C a region of fully inhibited tPRSS surrounding the source cells can be obtained in both conditions. In cell biology terms, this result suggests that even if tPRSS is secreted by a spatially different source cells or by multiple sources, the temporary inhibitor is still able to generate the first zone of fully inhibited tPRSS surrounding its source and the second zone of progressively active tPRSS away from it. An appropriated ratio [I]/[E] at the inhibitor source boundary is the necessary requirement for the existence of a physiologically meaningful extracellular zone of inhibited tPRSS activity. Interestingly, the ratio between SPINK2 and PRSS57, determined by proteomic analysis, is equal to 1.72 (Figure 2B), in agreement with the initial values used in the model.

DISCUSSION

Results reported here advance our knowledge on the biology and biochemistry of SPINK2, one member of the SPINK family of protease inhibitor, and shed new light on the physiological and pathophysiological meaning of its temporary enzymatic inhibition. Our work I) shows the physiological expression of SPINK2 in adult human

bone marrow HSPCs and in peripherally mobilized CD34[+] cells and provides a confirmation of its expression in a wide range of malignant hematological cancers,[11–13,43,44] II) defines, by *in vitro* experiments, the kinetic properties of trypsin inhibition by SPINK2 and SPINK2 degradation by trypsin and introduces them in mathematical relationships able to model long progress curve (TI-model), III) includes a thorough analysis of serine proteases gene expression in HSPCs of human adult bone marrow and shows for the first time the presence of PRSS2 (trypsin-2 or anionic trypsin) and PRSS57 (NSP4) in bone marrow HSPCs and mobilized CD34[+] hematopoietic cells, and IV) provides a quantitative theoretical background (TI-D theory) to the possible physiological meaning of the temporary enzyme inhibition properties of SPINKs.

The expression of SPINK2 in adult bone marrow HSPCs is in agreement with its expression in umbilical cord blood HSPCs[45] and the recent observation that SPINK2 is a marker of HSCs during human prenatal development.[46] Anjos-Afonso et al.[47] reported that a subpopulation from human cord blood, characterized by positivity for CD34 and EPCR (endothelial protein C receptor), represents the purest human HSC population. Indeed, analysis of their RNA-seq data (deposited at Gene Expression Omnibus GSE154263) confirmed the expression of SPINK2 in EPCR+ HSC in agreement with the results obtained in adult bone marrow. Moreover, by gene set enrichment,[47] EPCR+ cells were found to be enriched in a previously described gene module,[48] containing SPINK2 and Selectin L (SELL) as denominating genes.

We report here, for the first time, that SPINK2 is coexpressed in HSPCs with putative target serine protease, such as PRSS2 and PRSS57. The coexpression of members of the SPINK family and their target proteases is well established in several cell types. SPINK1 is coexpressed with trypsin (PRSS1 and PRSS2) in acinar cells of exocrine pancreas,[7] SPINK2 is coexpressed with the serine protease acrosin in spermatozoa,[9,10] and SPINK5 is coexpressed with kallikrein 5 and 7 in stratum granulosum keratinocytes.[49] Genetic defects of SPINKs are associated with the damage of the producing tissue, such as pancreatitis for SPINK1, azoo-oligospermia for SPINK2, and atopic dermatitis and erythrodermic ichthyosis for SPINK5.[4,9,10,50,51] These observations provide a strong argument in favor of the idea that the main role of SPINKs is the protection of the producing cells by an inappropriate early activation of the target proteases both intracellularly and extracellularly. Although it is reasonable that the protective role is an important aspect of SPINK functional role, less attention has been paid to the physiological meaning of the temporary enzymatic inhibition properties, well described few years after the discovery of the first member of the SPINK family.[1,26] Since the protective role might be played also by a persistent tight-binding inhibitor, the temporary inhibition exerted by SPINKs might be considered only a secondary phenomenon with little physiological relevance. However, the discovery of the expression of SPINK2 by HSPCs in a special tissue microenvironment, such as the bone marrow, prompted us to re-evaluate the possible role of temporary inhibition exerted by SPINKs. In this scenario the secretion of SPINK2 might block the tPRSS activity in the adjacent extracellular zone surrounding the SPINK-producing cells. In the same scenario the temporary inhibition exerted by SPINK2 might play a major part: it would provide a simple mechanism to reactivate the target proteases at a defined diffusion distance from the SPINK source cells. Therefore, a dual role can be ascribed to SPINK2 toward specific target proteases: an anti-proteolytic role in the vicinity of the producing cell and a proteolysis-permitting role away from it.

In this paper we provide a conceptual framework for this TI-D theory by applying the SDD model[39,40] to the production of SPINK2 by HSCs and its diffusion and degradation far from them. The SDD model is based on the second Fick's law plus a generic degradation term. The model generates a steady-state concentration gradient for a specific molecule diffusing from a source. In order to model the specific degradation of SPINK2 by its target PRSS, an additional degradation term was added, whose values are estimated by *in vitro* experiments with a model system composed of a modified porcine trypsin and recombinant SPINK2. The concentrations of the complex SPINK2/tPRSS and the enzymatic activities per unit volume at different distances from the source cell were calculated by the Morrison's quadratic equation for tight-binding inhibitors. Using diffusion constant values of $1–10\ \mu^2\ \text{sec}^{-1}$ and a SPINK2/tPRSS ratio of 2 at the source boundary, the size of the region of fully inhibited protease activity is 40–110 μm from the source, a width that is compatible with the hypothesis that this PRSS-inhibition zone might correspond to the HSCN. Of course, both parameters, D coefficient and pericellular ratio [I]/[E], need a direct experimental confirmation although data available for protein diffusion in different biological media are in agreement with the proposed D values[41,42] and transcriptomic and proteomic data of SPINK2/tPRSS ratios confirm the values adopted in our model.

Moreover, in the present work we identified two putative target serine proteases, PRSS2 and PRSS57, that are coexpressed and secreted with SPINK2 in HSPCs and mobilized CD34[+] cells. It should be kept in mind that SPINK2 action is not limited to target serine proteases coexpressed by the same producing cells but can be exerted also on tPRSS diffusing from a distant source. PRSS2 (anionic trypsin or trypsin-2) is one of the three isoforms of human trypsin and is produced as the proenzyme trypsinogen that should be proteolytically converted in active trypsin. The main enzyme involved in such activation in the small intestine is an enteropeptidase, a type II transmembrane serine protease 15 (TMPRSS15), localized to the brush border of the duodenal and jejunal mucosa. However, TMPRSS15 has not been reported to be expressed in bone marrow cell types, and it is likely that alternative mechanisms of trypsin activation do exist, such as autocatalytic activation[52] or activation by other not-yet-characterized membrane or secreted enzymes. PRSS57 (NSP4) is one of the four serine proteases expressed in differentiated neutrophils and the last to be discovered.[53,54] Here we report that PRSS57 is the only neutrophil serine protease whose expression is already high at the level of HSCs (Figure 5C). It has been shown that PRSS57 is activated intracellularly by cathepsin C (CTSC) and is secreted in active form.[53,54] A main difference between PRSS2 and PRSS57 is that PRSS2 expression seems to be mainly restricted to HSCs and HMPs while PRSS57 is highly expressed also in more specialized progenitors, such as GMPs, EPs, and MPs.

What could be the functional meaning of the predicted tPRSS-inhibition zone in the HSCN and the tPRSS-active zone away from it? A clue is provided from studies showing that anionic trypsin can efficiently activate the collagenolytic matrix metalloproteinase (MMP)-1, -8, and -13 and their activator MMP-3.[55] Trypsin-2 might be the initial serine protease activator of several other protease cascades involved in the remodeling of different extracellular matrix protein components and in the maintenance of the special fluid microenvironment of the bone marrow.[56] Moreover, anionic trypsin is an agonist of protease-activated membrane receptors (PAR2) triggering cell-specific responses[57] and may act as a local messenger involved in the anterograde communication between undifferentiated HSCs/HMP and the more committed progenitors surrounding them. On the other hand, the region of tPRSS inhibition could participate in the development of a more complex and stiffer extracellular matrix in the niche core, favoring cell adhesion and survival of HSCs in such microenvironment.[56] A role for the proteolytic breakdown of C-X-C motif chemokine 12 (CXCL12), vascular cell adhesion molecule 1 (VCAM-1), and stem cell factor (SCF) in the HSCN has been suggested for the mobilization of HSCs out of the bone marrow into the peripheral blood and serine proteases, and serine protease-activated MMPs might be involved in such process.[58–60]

An additional and different role can be suggested for the other putative tPRSS: PRSS57. Such enzyme is expressed both in the undifferentiated HSC/HMP and in more differentiate progenitor cells but with different SPINK2/PRSS57 expression ratios: 2.2–2.3 in undifferentiated HSC/HMP and 0.3–0.6 in more differentiated progenitors (GMP, MP, and EP). These data suggest that a region of fully inhibited PRSS57 could be present surrounding top-hierarchy HSC/HMP, while this inhibition zone is almost absent around more differentiated progenitors. In this scenario an increased secretion and diffusion of PRSS57 by GMP, MP, or EP could increase the PRSS57 concentration in the extracellular zone surrounding the HSC/HMP cells, thus decreasing the SPINK2/PRSS57 ratio around those cells. Therefore, PRSS57 could act as a retrograde signal moving from more differentiated progenitors toward less differentiated HSC/HMP. Indeed, it has been hypothesized that serine proteases released by G-CSF-activated GMPs might be responsible for the observed mobilization of CD34[+] cells from the HSCN to the blood.[58–60] PRSS57 might be the main serine protease involved in such process, and no study has previously investigated its action in bone marrow hematopoiesis.

Apart from diffusion, the TI-D model does not take into account other phenomena that might be involved in the protein mass movements in the extracellular fluid, such as the fluid exchanges between sinusoidal blood and bone marrow interstitial fluid. At the present stage of knowledge, it is extremely difficult to predict the influence of fluid flow dynamics on the diffusion of a soluble protein in the intercellular space. Indeed, the irregular geometry of arterial vessels and interconnecting sinusoidal capillaries,[61] the unclear architecture of perisinusoidal HSCNs,[38] the interplay of hydrostatic pressures in sinusoids and interstitium, and the lack of direct experimental determinations of these parameters make it difficult to introduce such variables in the model. On the other hand, interactions with components of the extracellular matrix, such as glycosaminoglycans and proteins, may strongly restrict the fluid flow-driven movement of secreted proteins by HSCs, making diffusion the main driving force for SPINK2 gradient generation. Another limitation of the model is the assumption that the protein secretion by source cells is constant and continuous, while a

pulsatile secretion or different forms of oscillatory secretion are possible. However, such limitations do not affect the main conclusion regarding the reactivation of the inhibited enzyme at a "physiologically mean-ingful distance" from the producing cell. Indeed, the demonstration that the value of $^{SPINK2}k_{cat}$ of a "model trypsin" is compatible with the reasonable anatomical size of the HSCN is an insight into the functional role of the temporary enzyme inhibition.[62]

On the basis of the above considerations, it is possible to make hypothesis on the role of the high expression of SPINK2 in chemoresistant leukemic cells. It has been suggested that adhesion of leukemic cells to a bone marrow vascular niche enhances their resistance to drug treatment, by promoting their entrance in a quiescent state that protects from drugs that act on quickly proliferating cells.[63] Therefore, SPINK2 is not only a marker of stemness for leukemic cells but could also directly participate to the formation of the bone marrow leukemic niche through its inhibitory action on serine proteases. In this sense SPINK2 might represent a new pharmacological target for specific subgroups of leukemias and other hematological malignancies.

Limitations of the study

It is worth mentioning that other physiological and pathophysiological actions of SPINKs, independent from the protease-inhibitory activity, have been suggested and investigated, particularly in the case of aberrant expression of SPINK1 in cancer.[64] These effects are not in contrast with the TI-D theory and could be integrated in the model enriching its functional complexity.

To the best of our knowledge, the TI-D theory is the first attempt to provide a theoretical background to the physiological role of temporary enzymatic inhibition in the context of a specific tissue and cell type. However, further experimental data are necessary to determine tissue-specific and enzyme-specific values for initial parameters of the TI-D model and to verify its predictions. Our analysis has shown that one important animal model, the mouse, is not suitable for SPINK2 experimental investigation since no expression of the mouse SPINK2 ortholog has been detected in the bone marrow. This difficulty can be overcome by the use of 3D tissue-engineered models of human bone marrow and by human hemato-poietic xenotransplatation in mice.[65–67] Moreover, the TI-D theory can be applied to other members of the SPINK family and to other tissue contexts, thus ensuring a wider range of experimental conditions where to test its general validity.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human cancer cell lines and subjects
- METHOD DETAILS
  - Single cell RNA-seq and proteomics profiling of human bone marrow hematopietic cells
  - Isolation of CD34$^+$
  - RNA extraction and qRT-PCR
  - Kinetics enzymatic studies
  - Western blotting
  - Temporary inhibition (TI)-model
  - Modelling of trypsin inhibitory action of SPINK2 by differential equations
  - The synthesis-diffusion-degradation (SDD) model
  - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.106949.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

D.F.C. and V.B. designed, planned, and coordinated the study and interpreted the data; D.F.C., V.B., and V.D.B wrote the manuscript; V.B., V.D.B. C.S., and A.P.P carried out CD34$^+$ purification, RNA extraction and qRT-PCR, protein extraction and western blot analysis; D.F.C., V.B., and V.D.B performed kinetic enzymatic experiments; D.F.C., V.B., V.D.B., and A.P.P. carried out bioinformatics analysis and prepared figures and tables; L.L.N. and P.B. performed samples collection and characterization of peripheral blood cells. All authors read and approved the final version of the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

REFERENCES

1. Kazal, L.A., Spicer, D.S., and Brahinsky, R.A. (1948). Isolation of a crystalline trypsin inhibitor-anticoagulant protein from pancreas. J. Am. Chem. Soc. 70, 3034–3040.

2. Whitcomb, D.C., Gorry, M.C., Preston, R.A., Furey, W., Sossenheimer, M.J., Ulrich, C.D., Martin, S.P., Gates, L.K., Jr., Amann, S.T., Toskes, P.P., et al. (1996). Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. Nat. Genet. 14, 141–145.

3. Chen, J.M., Mercier, B., Audrezet, M.P., and Ferec, C. (2000). Mutational analysis of the human pancreatic secretory trypsin inhibitor (PSTI) gene in hereditary and sporadic chronic pancreatitis. J. Med. Genet. 37, 67–69.

4. Witt, H., Luck, W., Hennies, H.C., Classen, M., Kage, A., Lass, U., Landt, O., and Becker, M. (2000). Mutations in the gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic pancreatitis. Nat. Genet. 25, 213–216.

5. Muller, N., Sarantitis, I., Rouanet, M., de Mestier, L., Halloran, C., Greenhalf, W., Férec, C., Masson, E., Ruszniewski, P., Lévy, P., et al. (2019). Natural history of SPINK1 germline mutation related-pancreatitis. EBioMedicine 48, 581–591.

6. Buchholz, I., Nagel, F., Klein, A., Wagh, P.R., Mahajan, U.M., Greinacher, A., Lerch, M.M., Mayerle, J., and Delcea, M. (2020). The impact of physiological stress conditions on protein structure and trypsin inhibition of serine protease inhibitor Kazal type 1 (SPINK1) and its N34S variant. Biochim. Biophys. Acta. Proteins Proteom. 1868, 140281.

7. Szabó, A., Toldi, V., Gazda, L.D., Demcsák, A., Tőzsér, J., and Sahin-Tóth, M. (2021). Defective binding of SPINK1 variants is an uncommon mechanism for impaired trypsin inhibition in chronic pancreatitis. J. Biol. Chem. 296, 100343.

8. Möritz, A., Lilja, H., and Fink, E. (1991). Molecular cloning and sequence analysis of the cDNA encoding the human acrosin-trypsin inhibitor (HUSI-II). FEBS Lett. 278, 127–130.

9. Lee, B., Park, I., Jin, S., Choi, H., Kwon, J.T., Kim, J., Jeong, J., Cho, B.N., Eddy, E.M., and Cho, C. (2011). Impaired spermatogenesis and fertility in mice carrying a mutation in the Spink2 gene expressed predominantly in testes. J. Biol. Chem. 286, 29108–29117.

10. Kherraf, Z.-E., Christou-Kent, M., Karaouzene, T., Amiri-Yekta, A., Martinez, G., Vargas, A.S., Lambert, E., Borel, C., Dorphin, B., Aknin-Seifer, I., et al. (2017). SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes. EMBO Mol. Med. 9, 1132–1149.

11. Barresi, V., Di Bella, V., Andriano, N., Privitera, A.P., Bonaccorso, P., La Rosa, M., Iachelli, V., Spampinato, G., Pulvirenti, G., Scuderi, C., et al. (2021). NUP-98 rearrangements led to the Identification of candidate biomarkers for primary induction failure in pediatric acute myeloid leukemia. Int. J. Mol. Sci. 22, 4575.

12. Xue, C., Zhang, J., Zhang, G., Xue, Y., Zhang, G., and Wu, X. (2019). Elevated SPINK2 gene expression is a predictor of poor prognosis in acute myeloid leukemia. Oncol. Lett. 18, 2877–2884.

13. Chen, T., Lee, T.-R., Liang, W.-G., Chang, W.-S.W., and Lyu, P.-C. (2009). Identification of trypsin-inhibitory site and structure determination of human SPINK2 serine proteinase inhibitor. Proteins 77, 209–219.

14. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science 347, 1260419.

15. Jardine, L., Webb, S., Goh, I., Quiroga Londoño, M., Reynolds, G., Mather, M., Olabi, B., Stephenson, E., Botting, R.A., Horsfall, D., et al. (2021). Blood and immune development in human fetal bone marrow and Down syndrome. Nature 598, 327–331.

16. Papatheodorou, I., Moreno, P., Manning, J., Fuentes, A.M., George, N., Fexova, S., Fonseca, N.A., Füllgrabe, A., Green, M., Huang, N., et al. (2020). Expression Atlas update: from tissues to single cells. Nucleic Acids Res. 48, D77–D83. https://doi.org/10.1093/nar/gkz947.

17. Ranzoni, A.M., Tangherloni, A., Berest, I., Riva, S.G., Myers, B., Strzelecka, P.M., Xu, J., Panada, E., Mohorianu, I., Zaugg, J.B., and Cvejic, A. (2021). Integrative single-cell RNA-seq and ATAC-seq analysis of human developmental hematopoiesis. Cell Stem Cell 28, 472–487.e7.

18. Kular, J., Tickner, J., Chim, S.M., and Xu, J. (2012). An overview of the regulation of bone remodelling at the cellular level. Clin. Biochem. 45, 863–873.

19. Setty, M., Kiseliovas, V., Levine, J., Gayoso, A., Mazutis, L., and Pe'er, D. (2019). Characterization of cell fate probabilities in single-cell data with Palantir. Nat. Biotechnol. 37, 451–460.

20. Nestorowa, S., Hamey, F.K., Pijuan Sala, B., Diamanti, E., Shepherd, M., Laurenti, E., Wilson, N.K., Kent, D.G., and Göttgens, B. (2016). A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood 128, e20–e31.

21. Reiss, K., Meyer-Hoffert, U., Fischer, J., Sperrhacke, M., Wu, Z., Dimitrieva, O., Krenek, P., Suchanova, S., Buryova, H., Brauer, R., and Sedlacek, R. (2011). Expression and regulation of murine SPINK12, a potential orthologue of human LEKTI2. Exp. Dermatol. 20, 905–910.

22. Hennrich, M.L., Romanov, N., Horn, P., Jaeger, S., Eckstein, V., Steeples, V., Ye, F., Ding, X., Poisa-Beiro, L., Lai, M.C., et al. (2018). Cell-specific proteome analyses of human bone marrow reveal molecular features of age-dependent functional decline. Nat. Commun. 9, 4004.

23. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589.

24. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. 50, D439–D444.

25. Morrison, J.F. (1969). Kinetics of the reversible inhibition of enzyme-catalysed reactions by tight-binding inhibitors. Biochim. Biophys. Acta 185, 269–286.

26. Laskowski, M., and Wu, F.C. (1953). Temporary inhibition of trypsin. J. Biol. Chem. 204, 797–805.

27. Walsh, R., Martin, E., and Darvesh, S. (2010). A method to describe enzyme-catalyzed reactions by combining steady state and time course enzyme kinetic parameters. Biochim. Biophys. Acta 1800, 1–5.

28. Bevc, S., Konc, J., Stojan, J., Hodošček, M., Penca, M., Praprotnik, M., and Janežič, D. (2011). ENZO: a web tool for derivation and evaluation of kinetic models of enzyme catalyzed reactions. PLoS One 6, e22265.

29. Nikolova, N., Tenekedjiev, K., and Kolev, K. (2008). Uses and misuses of progress curve analysis in enzyme kinetics. Cent. Eur. J. Biol. 3, 345–350.

30. Goličnik, M. (2011). Explicit analytic approximations for time-dependent solutions of the generalized integrated Michaelis-Menten equation. Anal. Biochem. 411, 303–305.

31. Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 46, D624–D632.

32. Raman, K., Trivedi, N.N., Raymond, W.W., Ganesan, R., Kirchhofer, D., Verghese, G.M., Craik, C.S., Schneider, E.L., Nimishakavi, S., and Caughey, G.H. (2013). Mutational tail loss is an evolutionary mechanism for liberating marapsins and other type I serine proteases from transmembrane anchors. J. Biol. Chem. 288, 10588–10598.

33. Chiravuri, M., Schmitz, T., Yardley, K., Underwood, R., Dayal, Y., and Huber, B.T. (1999). A novel apoptotic pathway in quiescent lymphocytes identified by inhibition of a post-proline cleaving aminodipeptidase: a candidate target protease, quiescent cell proline dipeptidase. J. Immunol. 163, 3092–3099.

34. Heissig, B., Hattori, K., Dias, S., Friedrich, M., Ferris, B., Hackett, N.R., Crystal, R.G., Besmer, P., Lyden, D., Moore, M.A.S., et al. (2002). Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of kit-ligand. Cell 109, 625–637.

35. Leu, S., Day, Y.-J., Sun, C.-K., and Yip, H.-K. (2016). tPA-MMP-9 Axis plays a pivotal role in mobilization of endothelial progenitor cells from bone marrow to circulation and ischemic region for angiogenesis. Stem Cells Int. 2016, 5417565.

36. Sillen, M., and Declerck, P.J. (2020). Targeting PAI-1 in cardiovascular disease: structural insights into PAI-1 functionality and inhibition. Front. Cardiovasc. Med. 7, 622473.

37. Crick, F. (1970). Diffusion in Embryogenesis. Nature 225, 420–422.

38. Comazzetto, S., Shen, B., and Morrison, S.J. (2021). Niches that regulate stem cells and hematopoiesis in adult bone marrow. Dev. Cell 56, 1848–1860.

39. Wartlick, O., Kicheva, A., and González-Gaitán, M. (2009). Morphogen gradient formation. Cold Spring Harb. Perspect. Biol. 1, a001255.

40. Fradin, C. (2017). On the importance of protein diffusion in biological systems: the example of the Bicoid morphogen gradient. Biochim. Biophys. Acta. Proteins Proteom. 1865, 1676–1686.

41. Peters, R. (1984). Nucleo-cytoplasmic flux and intracellular mobility in single hepatocytes measured by fluorescence microphotolysis. EMBO J. 3, 1831–1836.

42. Kumar, M., Mommer, M.S., and Sourjik, V. (2010). Mobility of cytoplasmic, membrane, and DNA-binding proteins in Escherichia coli. Biophys. J. 98, 552–559.

43. Hoefnagel, J.J., Dijkman, R., Basso, K., Jansen, P.M., Hallermann, C., Willemze, R., Tensen, C.P., and Vermeer, M.H. (2005). Distinct types of primary cutaneous large B-cell lymphoma identified by gene expression profiling. Blood 105, 3671–3678.

44. Elsayed, A.H., Rafiee, R., Cao, X., Raimondi, S., Downing, J.R., Ribeiro, R., Fan, Y., Gruber, T.A., Baker, S., Klco, J., et al. (2020). A six-gene leukemic stem cell score identifies high risk pediatric acute myeloid leukemia. Leukemia 34, 735–745.

45. He, X., Gonzalez, V., Tsang, A., Thompson, J., Tsang, T.C., and Harris, D.T. (2005). Differential gene expression profiling of CD34+ CD133+ umbilical cord blood hematopoietic stem progenitor cells. Stem Cells Dev. 14, 188–198.

46. Calvanese, V., Capellera-Garcia, S., Ma, F., Fares, I., Liebscher, S., Ng, E.S., Ekstrand, S., Aguadé-Gorgorió, J., Vavilina, A., Lefaudeux, D., et al. (2022). Mapping human haematopoietic stem cells from haemogenic endothelium to birth. Nature 604, 534–540.

47. Anjos-Afonso, F., Buettner, F., Mian, S.A., Rhys, H., Perez-Lloret, J., Garcia-Albornoz, M., Rastogi, N., Ariza-McNaughton, L., and Bonnet, D. (2022). Single cell analyses identify a highly regenerative and homogenous human CD34+ hematopoietic stem cell population. Nat. Commun. 13, 2048.

48. Velten, L., Haas, S.F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B.P., Hirche, C., Lutz, C., Buss, E.C., Nowak, D., et al. (2017). Human haematopoietic stem cell lineage commitment is a continuous process. Nat. Cell Biol. 19, 271–281.

49. Roelandt, T., Thys, B., Heughebaert, C., De Vroede, A., De Paepe, K., Roseeuw, D., Rombaut, B., and Hachem, J.P. (2009). LEKTI-1 in sickness and in health. Int. J. Cosmet. Sci. 31, 247–254.

50. Ohmuraya, M., Hirota, M., Araki, M., Mizushima, N., Matsui, M., Mizumoto, T., Haruna, K., Kume, S., Takeya, M., Ogawa, M., et al. (2005). Autophagic cell death of pancreatic acinar cells in serine protease

inhibitor Kazal type 3-deficient mice. Gastroenterology *129*, 696–705.

51. Hachem, J.-P., Wagberg, F., Schmuth, M., Crumrine, D., Lissens, W., Jayakumar, A., Houben, E., Mauro, T.M., Leonardsson, G., Brattsand, M., et al. (2006). Serine protease activity and residual LEKTI expression determine phenotype in Netherton syndrome. J. Invest. Dermatol. *126*, 1609–1621.

52. Kukor, Z., Tóth, M., and Sahin-Tóth, M. (2003). Human anionic trypsinogen: properties of autocatalytic activation and degradation and implications in pancreatic diseases. Eur. J. Biochem. *270*, 2047–2058.

53. Perera, N.C., Schilling, O., Kittel, H., Back, W., Kremmer, E., and Jenne, D.E. (2012). NSP4, an elastase-related protease in human neutrophils with arginine specificity. Proc. Natl. Acad. Sci. USA *109*, 6229–6234.

54. Perera, N.C., Wiesmüller, K.H., Larsen, M.T., Schacher, B., Eickholz, P., Borregaard, N., and Jenne, D.E. (2013). NSP4 is stored in azurophil granules and released by activated neutrophils as active endoprotease with restricted specificity. J. Immunol. *191*, 2700–2707.

55. Moilanen, M., Sorsa, T., Stenman, M., Nyberg, P., Lindy, O., Vesterinen, J., Paju, A., Konttinen, Y.T., Stenman, U.H., and Salo, T. (2003). Tumor-associated trypsinogen-2 (trypsinogen-2) activates procollagenases (MMP-1, -8, -13) and stromelysin-1 (MMP-3) and degrades type I collagen. Biochemistry *42*, 5414–5420.

56. Lee-Thedieck, C., Schertl, P., and Klein, G. (2022). The extracellular matrix of hematopoietic stem cell niches. Adv. Drug Deliv. Rev. *181*, 114069.

57. Zhao, P., Metcalf, M., and Bunnett, N.W. (2014). Biased signaling of protease-activated receptors. Front. Endocrinol. *5*, 67.

58. Winkler, I.G., Hendy, J., Coughlin, P., Horvath, A., and Lévesque, J.P. (2005). Serine protease inhibitors serpina1 and serpina3 are down-regulated in bone marrow during hematopoietic progenitor mobilization. J. Exp. Med. *201*, 1077–1088.

59. Tay, J., Levesque, J.-P., and Winkler, I.G. (2017). Cellular players of hematopoietic stem cell mobilization in the bone marrow niche. Int. J. Hematol. *105*, 129–140.

60. Hoggatt, J., Singh, P., Tate, T.A., Chou, B.K., Datari, S.R., Fukuda, S., Liu, L., Kharchenko, P.V., Schajnovitz, A., Baryawno, N., et al. (2018). Rapid mobilization reveals a highly engraftable hematopoietic stem cell. Cell *172*, 191–204.e10.

61. Bixel, M.G., Kusumbe, A.P., Ramasamy, S.K., Sivaraj, K.K., Butz, S., Vestweber, D., and Adams, R.H. (2017). Flow dynamics and HSPC homing in bone marrow Microvessels. Cell Rep. *18*, 1804–1816.

62. Wu, Q., Zhang, J., and Lucas, D. (2021). Anatomy of hematopoiesis and local microenvironments in the bone marrow. Where to? Front. Immunol. *12*, 768439.

63. Cogle, C.R., Bosse, R.C., Brewer, T., Migdady, Y., Shirzad, R., Kampen, K.R., and Saki, N. (2016). Acute myeloid leukemia in the vascular niche. Cancer Lett. *380*, 552–560.

64. Mehner, C., and Radisky, E.S. (2019). Bad tumors made worse: SPINK1. Front. Cell Dev. Biol. *7*, 10.

65. Abarrategi, A., Mian, S.A., Passaro, D., Rouault-Pierre, K., Grey, W., and Bonnet, D. (2018). Modeling the human bone marrow niche in mice: from host bone marrow engraftment to bioengineering approaches. J. Exp. Med. *215*, 729–743.

66. Chramiec, A., and Vunjak-Novakovic, G. (2019). Tissue engineered models of healthy and malignant human bone marrow. Adv. Drug Deliv. Rev. *140*, 78–92.

67. Salazar-Terreros, M.J., and Vernot, J.-P. (2022). In Vitro and in vivo modeling of normal and leukemic bone marrow niches: cellular senescence contribution to leukemia induction and progression. Int. J. Mol. Sci. *23*, 7350.

68. Barresi, V., Valenti, G., Spampinato, G., Musso, N., Castorina, S., Rizzarelli, E., and Condorelli, D.F. (2018). Transcriptome analysis reveals an altered expression profile of zinc transporters in colorectal cancer. J. Cell. Biochem. *119*, 9707–9719.

69. Chase, T., and Shaw, E. (1967). p-Nitrophenyl-p'-guanidinobenzoate HCl: a new active site titrant for trypsin. Biochem. Biophys. Res. Commun. *29*, 508–514.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-SPINK2	GeneTex	Cat#GTX32051
Rabbit anti-SPINK2	Sigma-Aldrich	Cat#HPA026813; RRID: AB_1857446
Mouse anti-GAPDH	Millipore	Cat#MAB374; RRID: AB_2107445
Rabbit anti-ACTB	Sigma-Aldrich	Cat. n. A2066; RRID: AB_476693
Biological samples		
Mobilized CD34$^+$ cells	This study	N/A
Chemicals, peptides, and recombinant proteins		
Recombinant human SPINK2	Sino Biological Inc.	Cat. n. 13636-H02H
Sequencing grade modified trypsin	Promega Corporation	Cat. n. V5111
4-nitrophenyl 4-guanidinobenzoate hydrochloride	Sigma-Aldrich	Cat. n. N8010
N-benzoyl-Phe-Val-Arg *p*-nitroanilide hydrochloride	Sigma-Aldrich	Cat. n. B7632
Critical commercial assays		
Dynabeads™ CD34 Positive Isolation Kit	Invitrogen	Cat. n. 11301D
RNeasy Mini Kit	Qiagen	Cat. n. 74104
High-Capacity RNA to cDNA Kit	ThermoFisher	Cat n. 4368813
SYBR® Green PCR Master Mix	ThermoFischer	Cat n. 4367659
Deposited data		
Human single cell RNAseq data	Setty et al.[19]	N/A
Mouse single cell RNA-seq data	Nestorowa et al.[20]	N/A
Cell lines RNA-seq data	Cell Model Passport	https://cellmodelpassports.sanger.ac.uk
Mass spectrometry proteomics data	Hennrich et al.[22]	N/A
Experimental models: Cell lines		
HCT-116	ATCC	CCL-247™
HL-60	ATCC	CCL-240
K562	ATCC	CCL-243
Kasumi-1	ATCC	CRL-2724
Jurkat	ATCC	TIB-152™
Oligonucleotides		
Primers for SPINK2, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A
Primers for PRSS1, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A
Primers for PRSS2, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A
Primers for PRSS57, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A
Primers for ELANE, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A
Primers for ACTB, see table in RNA extraction and qRT-PCR	Synthesized by https://eurofinsgenomics.eu	N/A

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Primers- BLAST	N/A	https://www.ncbi.nlm.nih.gov/tools/primer-blast/
ENZO web tool	Bevc S.[28]	http://enzo.cmm.ki.si/
EdgeR	N/A	https://bioconductor.org/packages/release/bioc/html/edgeR.html
Prism 8.0	GraphPad Software	https://www.graphpad.com/features

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Professor Daniele Filippo Condorelli (daniele.condorelli@unict.it).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Original samples and data are available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human cancer cell lines and subjects

Human colorectal adenocarcinoma cell line HCT116 (n.cat. CCL-247, male, adult), human promyeolocytic leukemia cell line HL-60 (n.cat. CCL-240, female 36-year-old), human acute chronic myelogenous leukemia cell line K562 (n. cat. CCL-243, female, 53-year-old), human acute myeloid leukemia cell line Kasumi-1 (n. cat. CRL-2724, male, 7 weeks) and human leukemia T-cell lymphoblastoid Jurkat cell line (n.cat. TIB-152, male, 14-year-old) were obtained from the American Type Culture Collection (ATCC). Karyotypes of HCT116 and Kasumi-1 cell lines were verified by SNP 6.0 array analysis (Affymetrix). Cell lines were routinely tested for mycoplasma using PCR-based approach. Cell lines were routinely maintained in RPMI medium (cat. No. 21875034 Gibco, Milan Italy) supplemented with 10% (v/v) foetal bovine serum (FBS), 1% of penicillin/streptomycin (cat. No. 15070063 Gibco, Milan Italy) and 1% of GlutaMax (cat. No. 35050061, Gibco, Milan Italy). Cells were cultured at 37°C in a humidified 5% CO_2 atmosphere.

Peripheral blood mononuclear cells were collected by leukapheresis from three human donors (3 males, 17-, 15- and 10-years-old), after treatment with G-CSF (10 μg/kg administered as subcutaneous injection once daily from at least 5 to 12 days) for a Peripheral Blood Stem Cell harvest.

METHOD DETAILS

Single cell RNA-seq and proteomics profiling of human bone marrow hematopietic cells

Single cell RNA-seq data, deposited from Setty et al.,[19] were downloaded from the "Human Cell Atlas Data Portal" (https://data.humancellatlas.org/explore/projects/091cf39b-01bc-42e5-9437-f419a66c8a45). UMAP graphs were elaborated through the Single Cell Expression Atlas (https://www.ebi.ac.uk/gxa/sc/home, accessed on May 2022). RNA-seq data of cancer cell lines were downloaded from Cell Model Passport (accessed on May 2022, https://cellmodelpassports.sanger.ac.uk/). The mass spectrometry proteomics data provided in Supplementary data 2 by Hennrich et al.[22] were used for analysis of SPINK2 and PRSS57 in HSPCs and MSCs.

Isolation of CD34⁺

CD34⁺ were isolated from granulocyte-colony stimulating factor (G-CSF)–mobilized peripheral blood stem cells. Mononuclear cells were first isolated from the blood samples by Ficoll density gradient using Ficoll-Paque Plus (Amersham). CD34⁺ cells were then purified using Dynabeads™ CD34 Positive Isolation Kit (Invitrogen), according to the manufacturer's recommendations: supernatant from washing steps was collected to obtain CD34⁻cells.

RNA extraction and qRT-PCR

RNA from CD34$^+$ cells and leukemia cell lines was extracted using Trizol reagent (Gibco, Milan, Italy) or commercial RNeasy Mini Kit (cat. no. 74104, Qiagen, Milan, Italy), respectively, and was quantified by the NanoDrop spectrophotometer.

Reverse transcription was performed starting from 1 μg of total RNA, using High-Capacity RNA to cDNA Kit (cat. No. 4368813 ThermoFisher, Milan, Italy). Primers were designed by the "Primers- BLAST" tool from NCBI (https://www.ncbi.nlm.nih.gov/tools/primer-blast/accessed on date 10 June 2020); in below table are reported for each transcript forward and reverse sequence primers, annealing temperature and fragment size. qRT-PCR analysis was performed using StepOneTM Real-Time PCR System by Applied Biosystems (Applied Biosystems, Foster City, CA, USA). Each sample was analyzed in triplicate and the average was normalized to human ACTB expression. Amplification conditions included a cycle at 95°C for 10 min followed by 40 cycles at 95°C for 15 s and 56–61°C for 1 min. As a negative control, reaction without cDNA was performed (no template control, NTC). The relative RNA expression level for each sample was calculated using the 2$^{-\Delta\Delta CT}$ method, as previously reported iesi et al.[68]

Primers details used for qRT-PCR

Transcript (Accession number)	Forward primer	Reverse primer	Annealing temperature	Fragment size
SPINK2 (NM_021114.4; NM_001271718.2)	2F TCCCTCAATTTGGTCTGTTTTC	3R CCTTCCCTGATTTTCATGCAC	60°C	151 bp
PRSS1 (NM_002769.5)	3F CTCACGTGCAGTAATCAACG	5R TCTTCTGGGCACAGCCATCA	56°C	329 bp
PRSS2 (NM_002770.4)	2F TTCTGGCTACCACTTCTGCG	3R TCCGGCTGTTGTATTTGGGG	58°C	182 bp
PRSS57 (NM_214710.5)	4F GGGCCACCTGACACTTACC	5R GAAGGAAACGAGGCCGTGA	56°C	124 bp
ELANE (NM_001972.4)	1F CGTGTCTTTTCCTCGCCTGT	2/3R CGTTTACATTCGCCACGCAG	57°C	207 bp
ACTB (NM_001101.5)	3F AGAGAGGCATCCTCACCCTG	4R ATAGCACAGCCTGGATAGCAA	89°C	240 bp

Kinetics enzymatic studies

Recombinant human SPINK2 was provided by Sino Biological Inc. (cat. No. 13636-H02H, Beijing, China). A DNA sequence encoding the human SPINK2 (NP_066937.1; Met1-Cys84) was expressed with the Fc region of human IgG1 at the carboxy-terminus in HEK293 cells and the mature protein (predicted N-terminal at Gln24) was purified at >90% as determined by SDS-PAGE. The recombinant protein consists of 299 amino acids and its predicted molecular mass is 33.6 kDa.

Porcine trypsin used in the present study is a sequencing grade modified trypsin (cat. No. V5111 Promega Corporation, Madison, WI, U.S.A.). Concentration of catalytically active enzyme was determined by active site titration using 4-nitrophenyl 4-guanidinobenzoate hydrochloride (cat. No. N8010, Sigma-Aldrich, Saint Louis, MO 63103, USA), as reported by Chase T Jr et al.[69]

Enzymatic reactions were determined by monitoring the conversion of N-benzoyl-Phe-Val-Arg p-nitroanilide hydrochloride (FVR-NA, cat. No. SIGMA N. B7632, Merck KGaA, Darmstadt, Germany) in the presence and absence of varying recombinant SPINK2 concentrations (3.75–120 nM) at 410 nm, using a BioTek Synergy H1 Multimode Reader (Agilent, Santa Clara, CA, U.S.A.).

Initial conversion rates were measured by using 0.125 mM FVR-NA, 3.2 nM trypsin and varying SPINK2 concentrations in 50 mM Tris-HCl pH 7.8, 1 mM CaCl2. Trypsin and SPINK2 were pre-incubated for 20 min and reaction was started by the addition of substrate. All assays were carried out at 30°C. In order to calculate Inhibition constant K_i the experimental data correlating initial reaction rates versus SPINK2 concentrations

were fitted to Morrison's quadratic equation for tight binding inhibitors by constraining total enzyme [E], initial substrate [S] and K_m to fixed values:

$$v = v_0 \left(1 - \frac{\left([E_T]+[I_T]+K_i\left(1+\frac{[S_T]}{K_m}\right) \right) - \sqrt{\left([E_T]+[I_T]+K_i\left(1+\frac{[S_T]}{K_m}\right) \right)^2 - 4[E_T][I_T]}}{2[E_T]} \right) \quad \text{(Equation 1)}$$

Where v is the reaction velocity in the presence of inhibitor, v_0 the reaction velocity in the absence of inhibitor, $[E_T]$ the total enzyme concentration, $[I_T]$ the total inhibitor concentration, $[S_T]$ the total substrate concentration, K_m the Michaelis and Menten constant, K_i the inhibition constant.

Western blotting

For immunoblotting, SPINK2 (60 nM) and trypsin (3.2 nM) were incubated at room temperature for 1, 3 and 6 h in a total volume of 100 μL. Then, 20 μL of each reaction containing 40 ng of SPINK2 were separated on 4–20% SDS-Polyacrylamide gel and transferred to nitrocellulose membranes. After blocking with Blotting-Grade Blocker (cat. No. 1706404, Biorad, Segrate, Milan, Italy) for 1 h at room temperature, membranes were incubated overnight at 4°C with the rabbit anti-SPINK2 (1:500, GeneTex cat. No. GTX32051). The day after, membranes were washed 3 times in TBS-T for 5 min and incubated for 1 h at room temperature with appropriate secondary antibody: goat anti-rabbit IgG-HRP (1:5000, Santa Cruz Biotechnology Inc.). Proteins bands were detected using ChemiDoc™ Imaging System (Biorad Segrate, Milan, Italy). Densitometry of western blot bands was performed using Image-J: after removing the background, the intensity of bands was measured generating the corresponding peaks and calculating the area for each one. The following procedure was applied on CD34$^+$ and CD34$^-$protein extracts: Mammalian Protein QProteome (QIAGEN, Milan, Italy) and BCA Protein Assay cat (Pierce™ BCA Protein Assay cat 23225) were used for extraction and determination of proteins concentration. 18μg for each protein sample were separated and transferred as already described. Membranes were incubated overnight at 4°C with the rabbit anti-SPINK2 (1:200, Sigma cat. No. HPA026813); the day after, membrane was stripped in Gly-HCl pH 1.8 buffer and incubated with mouse anti-GAPDH (1:5000, Millipore cat. No. MAB374) or rabbit anti-ACTB (1:5000, Sigma cat. No. A2066).

Temporary inhibition (TI)-model

Progress curve (substrate concentration vs time) of enzyme inhibition by competitive inhibitor I was modelled according to the following equation[27]:

$$[S]_t = [S]_0 \left(1 - \frac{V_{max}\left(\frac{[S]_0}{[S]_0+K_m - \left((K_m - K_{mapp})\frac{[I]}{[I]+K_i} \right)} \right)}{[S]_0} \right)^t \quad \text{(Equation 2)}$$

where $[S]_t$ is the concentration of substrate at time t, $[S]_0$ the initial substrate concentration, K_m and V_{max} represent the Michaelis-Menten constant and maximum reaction velocity in the absence of inhibitor, K_{mapp} represents the apparent K_m in the presence of the inhibitor at its concentration at time t, and Ki the inhibition constant. Progress curve of product vs time were obtained by setting $[P]_t$ (concentration of product formed at time t) equal to $[S]_0$-$[S]_t$.

In the case of a persistent inhibitor the concentration of the inhibitor [I] in Equation 2 is assumed to be constant during the time course. In the case of a Temporary Inhibitor (TI), i.e., SPINK2, [I] decreases with time according to the following relationship obtained by combining the first-order kinetic equation with the Michaelis and Menten equation applied to the interaction SPINK2/trypsin:

$$[I]_t = [I]_0 \left(1 - \frac{^{SPINK2}kcat[E_T]\left(\frac{[I]_0}{K_i+[I]_0} \right)}{[I]_0} \right)^t \quad \text{(Equation 3)}$$

Where $[I]_t$ is the concentration of SPINK2 at time t, $[I]_0$ is the initial concentration of SPINK2, $^{SPINK2}k_{cat}$ is the catalytic constant for SPINK2 degradation by trypsin, $[E_T]$ is the total concentration of trypsin (3.2 nM), K_i is the inhibition constant. The value of $[I]_t$ calculated by Equation 3 was substituted in Equation 2 for obtaining the time-course of product formation in the TI-model.

Modelling of trypsin inhibitory action of SPINK2 by differential equations

The web tool ENZO[28] was used for differential equation modelling and fitting. The seven differential equations were generated according to the scheme reported in Figure 4G. The following initial concentration of chemical species were used: [E]=3.2e-6 mM, [ES]=0, [EI]=0, [S]=125e-3 mM, [P]=0, [I]=60e-6 mM, [D]=0. The following fixed values were used for the kinetic constants: k0 = 1751 mM^{-1} min^{-1}, k1 = 1.2 x 10^{-6} min^{-1}, k2 4375 min^{-1}, obtained by initial velocity assays with modified porcine trypsin and FVR-NA substrate, and k3 = 1.8 x 10^5 mM^{-1} min^{-1} and k4 = 3.29 x 10^{-3} min^{-1}, using values previously reported by Szabo et al.[7] for SPINK1. k5 constant was obtained by fitting to the experimental progress curve of product formation.

The synthesis-diffusion-degradation (SDD) model

The SDD model[39,40] was used to generate steady-state concentration gradients of SPINK2 secreted by HSPCs in the hematopoietic niche. Moreover, the same model was used to simulate concentration gradients of tPRSS co-secreted by HSPCs or produced by a different distant cell type. Since gradients are often symmetric around the source the theoretical analysis has been performed for one dimension in a diffusion path of length L. The following diffusion equation with a linear degradation term was used:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} - k_c \qquad \text{(Equation 4)}$$

Where c is the concentration of the diffusing protein (tPRSS or SPINK2), D the diffusion coefficient, k the rate constant of degradation (sec^{-1}), t is time and x the distance form source. Two boundary conditions must be added to Equation 4, $D\frac{\partial c}{\partial x}|_{x=0} = -j_0$ to account for the influx at the source (x=0), and $D\frac{\partial c}{\partial x}|_{x=L} = 0$ at the end of the diffusion path (x=L).

The solution of Equation 4 at the steady-state condition (concentration gradients do not change with time) is the following:

$$c_x = c_0 e^{-x/\lambda} \qquad \text{(Equation 5)}$$

where c_x is the protein concentration at distance x, c_0 is the gradient amplitude, i.e. the concentration at the source boundary (x = 0), and λ is the decay length. c_0 and λ are represented by the following relationships:

$$c_0 = \frac{j_0}{\sqrt{Dk}} \text{ and } \lambda = \sqrt{\frac{D}{k}}$$

where J_0 is the flux of molecules across the source boundary expressed as μmoles/(μ^2 x sec).

In some simulations an inverse gradient of tPRSS was used in order to simulate the production of tPRSS from a cell source different of that of SPINK2. Equation 5 was modified in the following way to obtain an increasing gradient of tPRSS starting from SPINK2 source boundary:

$$c_x = c_0 e^{x/\lambda}$$

Where c_x is the concentration of tPRSS at distance x from SPINK2 source and c_0 is the concentration of tPRSS at the SPINK2 source boundary (x=0). In other simulations, as indicated in the result section, c_x of tPRSS was considered constant.

Since we aimed to compare a persistent protease-inhibitor (not degraded by its target protease) with a temporary one (degraded by its target protease), the rate constant of degradation k was considered as the sum of two terms: k= k' + k''. k' is related to a general linear protein degradation process, is not dependent by c_0, is assumed to be equal for tPRSS and SPINK2, is constant in the considered diffusion path L, and is negligible in comparison to k'' ($k'' \gg k'$). k'' is related to the specific degradation of SPINK2 by tPRSS (temporary inhibition phenomenon) and it is set equal to zero in the case of a putative persistent protease inhibitor.

The comparison between the persistent inhibition model and the temporary inhibition one is performed by fixing an equal value of c_0 and D for the two conditions and by calculating (Equation 5) the concentrations of inhibitor and protease at different distance x from source (c_x). As reported in the result section, degradation of SPINK2 by trypsin is not a linear process and k'' values change as a function of the initial concentration of SPINK2. Therefore, addition of k'' introduces a non-linear component in Equation 4. However, a different value of k'' was calculated and used for each c_0 tested ($k''=f(c_0)$) and the temporary inhibition was approximated by a linear process determining an exponential decay of the inhibitor (see Figure 8 in the result section).

In order to calculate the concentration of the complex $[EI]_x$ formed by the target PRSS and the SPINK2 inhibitor, at different distance x from the source and in the presence of a fixed concentration of the competing physiological substrate S for the enzyme, the following Morrison quadratic equation was used:

$$[EI]_x = \frac{\left([E_T]_x + [I_T]_x + K_d\left(1 + \frac{[S]}{K_m}\right)\right) - \sqrt{\left([E_T]_x + [I_T]_x + K_d\left(1 + \frac{[S]}{K_m}\right)\right)^2 - 4[E_T]_x[I_T]_x}}{2} \quad \text{(Equation 6)}$$

Where $[EI]_x$ is the concentration of the complex Enzyme/Inhibitor at distance x, $[E_T]_x$ is the total concentration (µM) of the target protease enzyme at x, $[I_T]_x$ (µM) is the total concentration of the inhibitor at x, Kd is the dissociation constant of the complex [EI], [S] is the concentration of the protein substrate of the target protease and K_m is the Michaelis and Menten constant of the target protease. The Kd value experimentally obtained by Szabo et al.[7] for SPINK1 and the Kd values determined in previous section for SPINK2 were tested in such simulations. The [S] is assumed constant in all distances x and a value of 125 µM was used for simulations. The K_m value of modified porcine trypsin experimentally determined for the FVR-NA substrate (2450 µM) was used for simulations.

In order to obtain an estimate of enzyme activity per unit volume at different distances from the producing cell, the velocities of conversion of protein substrate (S_v), expressed as uM/min, were calculated at different distances x from the source with the following formula:

$$^S v_x = \frac{^S kcat[E_T]_x[S_T]_x}{K_m + [S_T]_x}\left(1 - \frac{[E_I]_x}{[E_T]_x}\right) \quad \text{(Equation 7)}$$

Statistical analysis

Publically available normalized CPM data in Figures 1C, 1D, 5A–5C and 5E–5G are reported as means. The number of analyzed cells (n) for each population is indicated in the legend of Figure 1.

Publically available TPM data in Figure 3A are reported as means ± SD; the number of tissue samples for each category (n) is indicated in the graph of Figure 3A. Data in Figures 2A, 3E and 5D are expressed as Fold-Changes in comparison to control cells; results are reported as means of triplicate assays. Data in Figure 4A are means ± SEM of two independent experiments. All statistical analysis were performed with GraphPad Prism 8 software (v.8; GraphPad Software Inc., La Jolla, CA, USA).