Genetically determining individualized clinical reference ranges for the biomarker tryptase can limit unnecessary procedures and unmask myeloid neoplasms

Authors: Jack Chovanec¹, Ilker Tunc², Jason Hughes³, Joseph Halstead⁴, Allyson Mateja⁵, Yihui Liu¹, Michael P. O’Connell¹, Jiwon Kim¹, Young Hwan Park¹, Qinlu Wang⁶, Quang Le⁷, Mehdi Pirooznia², Neil N. Trivedi⁸, Yun Bai⁹, Yuzhi Yin⁹, Amy P. Hsu¹⁰, Josh McElwee¹¹, Sheryce Lassiter¹², Celeste Nelson¹, Judy Bando¹, Thomas DiMaggio¹³, Julij Šelb¹⁴, Matija Rijavec¹⁴, Melody C. Carter⁹, Hirsh D. Komarow⁹, Vito Sabato¹⁵, Joshua Steinberg¹⁶, Kurt M. Hafer¹⁷, Elizabeth Feuille¹⁸, Christopher S. Hourigan¹⁹, Justin Lack²⁰, Paneez Khoury²¹, Irina Maric²², Roberta Zanotti²³, Patrizia Bonadonna²⁴, Lawrence B. Schwartz⁷, Joshua D. Milner²⁵, Sarah C. Glover²⁶, Didier G. Ebo¹⁵, Peter Korošec¹⁴, George H. Caughey⁸, Erica H. Brittain²⁷, Ben Busby²⁸, Dean D. Metcalfe⁹, and Jonathan J. Lyons.¹†

Affiliations: ¹Translational Allergic Immunopathology Unit, Laboratory of Allergic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ²Bioinformatics and Computational Biology Core, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; ³Foundation Medicine, Cambridge, MA; ⁴Birmingham Women's and Children's NHS Foundation Trust, Birmingham, England; ⁵Clinical Monitoring Research Program Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD; ⁶Bioinformatics and Computational Biosciences Branch, Office of Cyber Infrastructure and Computational Biology, NIAID, NIH, Bethesda, MD; ⁷Department of Internal Medicine, Virginia Commonwealth University, Richmond, Virginia; ⁸Cardiovascular Research Institute and Department of Medicine, University of California San Francisco, San Francisco, CA,
and Veterans Affairs Medical Center, San Francisco, CA; ²⁹Mast Cell Biology Section, Laboratory of Allergic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ³⁰Immunopathogenesis Section, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ³¹Nimbus Therapeutics, Cambridge, MA; ³²Clinical Research Directorate, Frederick National Laboratory for Cancer Research, Frederick, MD; ³³Fungal Pathogenesis Section, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ³⁴University Clinic of Respiratory and Allergic Diseases, Golnik, Slovenia; ³⁵Department of Immunology, Allergology, and Rheumatology, University of Antwerp and Antwerp University Hospital, Infla-Med, Centre of excellence, University of Antwerp, Antwerp, Belgium; ³⁶Division of Allergy and Clinical Immunology, Department of Pediatrics, Medical College of Wisconsin, Milwaukee, WI; ³⁷Department of Medicine, Stanford University, Palo Alto, CA; ³⁸Division of Allergy and Clinical Immunology, Department of Pediatrics, Weill Cornell Medical College, New York, NY; ³⁹Myeloid Malignancies Section, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD; ⁴⁰NIAID Collaborative Bioinformatics Resource (NCBR), National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ⁴¹Human Eosinophil Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD; ⁴²Hematology Service, Department of Laboratory Medicine, Clinical Center, National Institutes of Health, Bethesda, MD; ⁴³Department of Medicine, Section of Hematology, Verona University Hospital, Verona, Italy; ⁴⁴Allergy Unit, Verona University Hospital, Verona, Italy; ⁴⁵Division of Allergy, Immunology and Rheumatology, Columbia University, New York, NY; ⁴⁶Division of Digestive Diseases, Department of Medicine,
University of Mississippi Medical Center, Jackson, MS;  
27Biostatistics Research Branch, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD;  
28National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD

†Correspondence: Jonathan J. Lyons, M.D.  
9000 Rockville Pike,  
Building 29B, Room 5NN18, MSC 1889  
Bethesda, MD 20892  
Phone: (301) 443-5250  
Fax: (301) 480-5757  
Email: jonathan.lyons@nih.gov

Funding: This project has been funded in whole or in part with federal funds from the Division of Intramural Research of the National Institute of Allergy and Infectious Diseases, NIH and with federal funds from the National Cancer Institute, National Institutes of Health, under Contract Nos. HHSN26120150003I and 75N910D00024. This project was also funded in part with federal funds from the Division of Intramural Research of the National Heart, Lung, and Blood Institute, NIH. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The involvement of BB and JH was supported by the Intramural Research Program of the National Library of Medicine. SCG is supported by US Public Health Service Grant NIH R21DE028378. VS is a Senior Clinical
Researcher of the Research Foundation Flanders/ Fonds Wetenschappelijk Onderzoek (FWO: 1804518N). DE is a Senior Clinical Researcher of the Research Foundation Flanders/Fonds Wetenschappelijk Onderzoek (FWO: 1800614N).

Conflict of interest statement: VCU receives royalties from Thermo Fisher for their tryptase test that are shared with LBS as its inventor. None of the remaining authors have relevant conflicts of interest to report.
ABSTRACT

Serum tryptase is a biomarker used to aid in the identification of certain myeloid neoplasms, most notably systemic mastocytosis, where baseline (BST) levels >20 ng/mL are a minor criterion for diagnosis. Whereas clonal myeloid neoplasms are rare, the common cause for elevated BST is the genetic trait hereditary alpha-tryptasemia (HαT) caused by increased germline TPSAB1 copy number. To date, the precise structural variation and mechanism(s) underlying elevated BST in HαT and the general clinical utility of tryptase genotyping, remain undefined. Through cloning, long-read sequencing, and assembling of the human tryptase locus from an individual with HαT, and validating our findings in vitro and in silico, we demonstrate that BST elevations arise from over-expression of replicated TPSAB1 loci encoding wild-type α-tryptase due to co-inheritance of a linked over-active promoter element. Modeling BST levels based upon TPSAB1 replication number we generate new individualized clinical reference values for the upper limit of 'normal'. Using this personalized laboratory medicine approach, we demonstrate the clinical utility of tryptase genotyping, finding that in the absence of HαT, BST levels >11.4 ng/mL frequently identify indolent clonal mast cell disease. Moreover, substantial BST elevations (e.g., >100 ng/mL) which would ordinarily prompt bone marrow biopsy, can result from TPSAB1 replications alone and thus be within 'normal' limits for certain individuals with HαT.
INTRODUCTION

The tryptase locus is structurally complex in humans (Table 1). Approximately two-thirds of people have α-tryptase, encoded at TPSAB1 on one or both alleles, whereas everyone has β-tryptase encoded at one or both TPSB2 alleles, as well as at non-α-tryptase encoding TPSAB1 loci (1-3). Copy number gain and loss of tryptase gene sequences encoding both α- and β-tryptases have also been reported (4). However, the structures of such copy number variants (CNV) remain unknown. The most common CNVs observed among Western populations are TPSAB1 gene replications encoding α-tryptase, a genetic trait known as hereditary alpha-tryptasemia (HαT) (5).

HαT is inherited in an autosomal dominant pattern and has been shown to affect nearly 6% of the general population in the U.S., U.K., and E.U (6-9). De novo replications have not been reported. TPSAB1 replications are associated with elevated basal serum tryptase (BST) levels of at least 2-3-fold higher than median BST levels of healthy individuals without HαT (10). While it has been hypothesized that additional TPSAB1 copies identified in HαT are present within the tryptase locus at chromosome 16p13.3, and that over-expression of α-tryptase yields the disproportionate increases in BST levels seen relative to copy number (5), this has never been demonstrated. Moreover, since only a small region of TPSAB1 has been probed to demonstrate gene replication (5, 11), it remains unknown whether extra-allelic copies of TPSAB1 encode wild-type or novel sequences containing homology with α-tryptase at the probe site.

Whereas HαT is the most common heritable cause for elevated BST, the clonal mast cell disorder systemic mastocytosis (SM) is a common acquired cause (10, 12), and BST levels are used routinely as a biomarker to screen for this disorder. Currently, the WHO criteria define a BST level >20 ng/mL as a minor criterion for the clinical diagnosis of SM (13) and this cut-off is frequently used in clinical decision making as an indication for bone marrow biopsy in
symptomatic individuals (14). However, approximately one out of every four individuals with H\(\alpha\)T has a BST level >20 ng/mL, representing an estimated 7.5 million people in the U.S. alone (4, 15). Because H\(\alpha\)T has been shown to augment immediate hypersensitivity symptoms in a number of conditions (7, 8), many of these individuals are likely to undergo unwarranted invasive work-up for SM, including bone marrow biopsy, if tryptase genotype is not taken into account. Conversely, because H\(\alpha\)T appears to account for most elevations in BST, observed elevations in patients who do not have H\(\alpha\)T could be considerably enriched for other pathologies - most notably myeloid neoplasms - warranting workup when such elevations in BST may be modest (i.e., <20 ng/mL).

To address these questions, we cloned, long-read sequenced, and assembled the human tryptase locus containing a \(TPSABI\) replication, finding the coding sequence to be wild-type \(\alpha\)-tryptase. However, a series of unique proximal non-coding variants were also identified that distinguished replicated (\(\alpha^{DUP}\)-tryptase) from non-replicated (\(\alpha^{WT}\)-tryptase) sequences. An expanded DNA motif within the 5'-UTR was also linked to \(TPSABI\) replication-associated variants and demonstrated increased \textit{in vitro} promoter activity relative to the paralogous region in the non-replicated promoter. Using \textit{in vitro} and \textit{in silico} analyses of RNA sequences, we confirmed the relative over-expression of \(\alpha^{DUP}\)-tryptase sequences in primary basophils, cultured mast cells, and publicly available RNA-sequence datasets. Applying this knowledge, we generated a new genetic-based model for BST clinical reference ranges based upon \(TPSABI\) replication number and examined the potential real-world impact of coupling tryptase genotyping with these newly defined values in the work-up of patients with clonal mast cell disorders. The potentially practice-changing implications for identification of indolent clonal mast cell disease, the diagnosis of SM, and the elimination of unnecessary bone marrow biopsies are discussed herein.
RESULTS

Increased TPSAB1 copy number occurs at the tryptase locus and encodes wild-type alpha-tryptase

A bacterial artificial chromosome (BAC) library was generated from an individual homozygous for TPSAB1 duplications (genotype αα/β:αα/β) and a clone containing the complete human tryptase locus was identified. Using single molecule real time (SMRT) technology, we sequenced and subsequently assembled the locus de novo (Fig. 1a), allowing study of the structure and sequence of this variant-containing allele for the first time. As anticipated, the assembly displayed marked structural dissimilarity with the reference human genome (GRCh37/hg19) but did not contain novel sequence (Fig. S1, and online data repository). The additional TPSAB1 copy encoding α-tryptase (α\text{DUP}) was located within the tryptase locus mapping to 16p13.3 between TPSB2 and the non-replicated TPSAB1 locus, as had been suggested based upon prior in vitro experiments (5). The coding sequence was identical to that of the adjacent wild-type α-tryptase sequence (α\text{WT}) (Table S4; online repository), and both sequences were in reverse orientation.

The sequence at TPSB2 in our BAC clone was identified as β1-tryptase. While β1-tryptase at TPSB2 is not common, it has been reported in approximately 5% of individuals screened from the HapMap cohort (16). Given that this prevalence was comparable to the prevalence of HαT reported in Western populations, we obtained the samples reported to contain β1-tryptase at TPSB2 in order to determine whether they were all individuals with HαT. However, only 4 out of 15 of these individuals (27%) with β1-tryptase at TPSB2 were found to have HαT. Thus, β1-tryptase at TPSB2 is not exclusive to HαT.

TPSAB1 replications encoding α\text{DUP}-tryptase are linked to an expanded promoter
Elevated BST levels among patients with HαT have been shown to result from increased constitutive release of pro-tryptases (5, 17). To determine whether this increase results from over-expression of αDUP-tryptase transcripts, we compared non-coding sequences at the two TPSAB1 loci and identified an expanded (~800 bp) repetitive DNA motif approximately 1 kilobase (kB) 5’ of the consensus transcription start site for αDUP-tryptase (Fig. 1b). This sequence was approximately 700 bp longer than the paralogous region of the αWT promoter containing 27 repeated and 3 repeated sequences, respectively (Fig. 1c). Repeat regions at the same 5’ location were also observed in publicly available assemblies of αWT sequences and confirmed by gel electrophoresis in amplified αWT-tryptase and αDUP-tryptase promoters from 9 unrelated individuals with HαT (Fig. S2). Subsequent cloning and Sanger sequencing of three additional αWT and αDUP-tryptase promoters confirmed these relative differences (Fig. S3). Together these data demonstrate an expanded DNA repeat motif in the promoter of TPSAB1 replications encoding αDUP-tryptase that is consistently co-inherited.

**Increased basal αDUP-tryptase expression is associated with elevated basal promoter activity**

To quantify the functional outcome of the expanded promoter region linked to αDUP-tryptase sequences we next interrogated total α- and β-tryptase as well as αWT- and αDUP-tryptase transcript expression arising from TPSAB1 and TPSB2. Cultured primary human mast cells and isolated basophils from individuals with HαT expressed higher levels of total α-tryptase transcripts, and over-expressed αDUP-tryptase sequences relative to other tryptase isoforms (Fig. 1d,e). Importantly, total α-tryptase gene expression levels were several orders of magnitude higher than those of β-tryptases, even when they were present in allelic balance (e.g., 2α,2β versus 3α,3β).
To determine if the expanded 5' sequences uniquely associated with α<sup>DUP</sup>-tryptase contributed to the relative over-expression of α<sup>DUP</sup>- to α<sup>WT</sup>-tryptase, we next cloned the promoters containing these paralogous regions into fluorescent reporter plasmids using our BAC assembly and other SMRT-based de novo assemblies of the locus as sequence references (see methods for referenced assemblies). These constructs were sequence-verifed and transfected into MonoMac-6 cells – an AML line that expresses tryptases (18, 19). The expanded α<sup>DUP</sup>-tryptase promoter demonstrated significantly greater basal activity relative to the α<sup>WT</sup>-tryptase promoter (Fig. 2a).

This in vitro finding correlated directly with BST levels among individuals with HαT where the absolute number of TPSAB1 replications was found to best correlate with BST level regardless of tryptase genotype derived from TPSAB1 and TPSB2 (Fig. 2b). Therefore, elevated BST in HαT results from over-expression of α<sup>DUP</sup>-tryptase at replicated TPSAB1 loci linked to a promoter with increased basal activity.

A unique haplotype demonstrates conservation of the identified over-active promoter at replicated TPSAB1 loci

To examine the generalizability of our finding of an expanded promoter linked to α<sup>DUP</sup>-tryptase at TPSAB1, we again examined proximal non-coding sequences associated with different tryptase isoforms and identified five unique 5' variants – two substitutions in intron 1, two in the 5' UTR, and one in the proximal promoter (Fig. 1b). Using these variants, we defined haplotypes, that when combined with coding sequences, uniquely distinguished α<sup>DUP</sup>-tryptase (CACCT) from α<sup>WT</sup>-tryptase (GGTTC), β1-tryptase (GACCC), and δ-tryptase (GGTTC) in our BAC assembly (Fig. 1b; see Table S1 for complete contextual sequences).
Next, we examined publicly available genome assemblies to determine if these haplotypes were conserved. Critically, TPSAB1 and TPSB2 have been shown to be in nearly complete linkage disequilibrium and are co-inherited in virtually all individuals such that major haplotypes exist and inheritance can be inferred (16, 20). The major tryptase isoforms encoded at TPSAB1 are \( \beta_1 \)-tryptase followed by \( \alpha^{WT} \)-tryptase, and at TPSB2 are \( \beta_3 \)-tryptase and \( \beta_2 \)-tryptase, when on \( \beta/\beta \) and \( \alpha/\beta \) alleles respectively. As expected we observed GACCC linked to \( \beta_1 \)-tryptase at TPSB2 on \( \beta_1/\beta_3 \) (\( \beta/\beta \)) alleles, and GGTTC linked to \( \alpha^{WT} \)-tryptase sequences at TPSAB1 on \( \alpha^{WT}/\beta_2 \) (\( \alpha/\beta \)) alleles in all 13 queried assemblies as well as well as in the GRCh37/hg19 reference assembly (Table 1). Additional isoform-specific haplotypes at these two loci - GACCT linked to \( \beta_2 \)-tryptase at the TPSB2 locus only when \( \alpha^{WT}/\beta_2 \) (\( \alpha/\beta \)) alleles were present and CACCT linked to \( \beta_3 \)-tryptases at the TPSB2 locus only when \( \beta_1/\beta_3 \) (\( \beta/\beta \)) alleles were present - were also identified. Neither the \( \alpha^{DUP} \)-tryptase-specific promoter expansion nor the linked CACCT haplotype were observed in association with \( \alpha^{WT} \)-tryptase sequences in any of the seven assemblies containing \( \alpha^{WT}/\beta_2 \) (\( \alpha/\beta \)) (Table 1).

Using these haplotypes, we next re-examined genome sequence reads from 183 individuals without increased TPSAB1 copy number, and from 32 affected individuals from 10 families with H\( \alpha \)T (5). Employing a paired-read approach, we found that GACCC (\( \beta_1 \)-tryptase) and CACCT (\( \beta_3 \)-tryptase) were present in association with 98\% (324/331) of individuals with \( \beta/\beta \) alleles. Moreover, the CACCT haplotype consistently segregated with the \( \beta/\beta \) haplotype in the 16 families without H\( \alpha \)T in whom tryptase sequence inheritance could be examined. Whereas all \( \alpha/\beta \) haplotypes (\( n = 6 \)) were associated with the \( \alpha^{WT} \)-tryptase-linked GGTTC and \( \beta_2 \)-tryptase-linked GACCT (Table 2), the CACCT haplotype linked to \( \alpha^{DUP} \)-tryptase in our BAC clone, segregated...
universally with \textit{TPSAB1} replications in all 32 individuals with H\textalpha T. The GACCC haplotype – linked to \(\beta 1\)-tryptase most often present at \textit{TPSAB1} but present at \textit{TPSB2} in our BAC clone – was also found to be universally present among individuals with H\textalpha T and co-segregated with \(\alpha^{\text{DUP}}\)-tryptase containing alleles suggesting that \(\beta 1\)-tryptase at \textit{TPSB2} is frequently present on alleles containing \textit{TPSAB1} replications, as seen in our BAC clone (Table 2). Therefore, these \textit{in silico} findings suggest conservation of the unique promoter haplotypes we identified in our BAC assembly, and that over-expression of \(\alpha^{\text{DUP}}\)-tryptase could be a generalizable phenomenon.

To confirm these findings at scale \textit{in vitro}, we had to overcome the repetitive and GC-rich nature of the region. To do so we developed a ddPCR assay capable of detecting the most proximal variant which distinguished the \(\alpha^{\text{DUP}}\)-tryptase promoter haplotype (CACCT) from that of \(\alpha^{\text{WT}}\)-tryptase (GGTTC) (see Fig. S4 for description and representative data). The corresponding variant listed in the reference genome (GRCh37/hg19 16:1290818C>T) has only been reported in association with \textit{TPSB2} – ostensibly linked to dominant isoform \(\beta 3\)-tryptase at this locus. However, the reverse primer in our assay was designed to only hybridize with \(\alpha\)-tryptase sequences.

The CACCT variant was confirmed to be universally present in linkage with \(\alpha^{\text{DUP}}\)-tryptase in 120 affected individuals with increased \textit{TPSAB1} copy number from 101 families. Conversely, CACCT was never found in linkage with \(\alpha^{\text{WT}}\)-tryptase in 81 unaffected individuals from 69 families who did not have increased \textit{TPSAB1} copy number but carried one or two \(\alpha^{\text{WT}}\)-tryptase-encoding copies (Fig. 2c). Furthermore, the ratios of identified promoter copy number confirmed that all 166 H\textalpha T-associated \(\alpha^{\text{DUP}}\)-tryptase sequences on 156 alleles were linked to the C>T variant (Fig. 2d) and by inference, the expanded promoter.
Interestingly, in 30/120 individuals with HαT, αWT-tryptase was not linked to GGTTCC; 10 of these αWT-tryptase promoters appeared to contain the C>T missense, and the remaining 19 alleles may have had this or another unknown missense variant or indel, the latter of which may have resulted in failed probe or primer hybridization. Importantly, our in silico analysis identified a GGTTT haplotype linked to αWT-tryptase in one family with HαT. In a second family a single base-pair deletion in the -6 position from C>T was observed, which we have labelled GGTT(-6delC)C was also identified (Table 2). This deletion interfered with probe hybridization in our assay potentially accounting for some of the 19 αWT-tryptase containing alleles with ambiguous promoter linkage. The αWT-tryptase promoter from these samples also failed to amplify and was not successfully cloned, suggesting yet another possible αWT-tryptase promoter may exist. Indeed, public database interrogation yielded at least six different α-tryptase sequences exhibiting a high degree of homology (Table S4). Regardless, αWT-tryptase containing alleles with undefined promoter haplotypes, or even those linked to the C>T intronic variant, were not associated BST differences in vivo (Fig. S5), indicating that these undefined promoters are not associated with differences in gene expression at linked αWT-tryptase encoding TPSAB1 loci. Collectively, these data support conservation of a unique haplotype linked to an over-active promoter element at replicated TPSAB1 loci that is associated with over-expression of α-tryptase by mast cells and basophils.

**αDUP-Tryptase transcripts are over-expressed in public datasets**

Using α- and β-tryptase isoform-specific coding variants and the uniquely linked proximal non-coding haplotypes, a 39-base consensus sequence was defined that contained four variants which could distinguish αWT-, αDUP-, and β-tryptase transcripts; given the proximity of the first variant
to the transcriptional start site, only the distal three variants were used for read alignment (Fig. 1b, Table S2). Querying 58 non-disease-associated transcriptome datasets, 863/4160 samples were identified with exact transcript matches to ≥1 of the three defined 39-bp consensus sequences. αDUP-Tryptase was observed in 3% (25/863) and αWT-tryptase transcripts were identified in 34% of individuals (289/863) (Table S5). Because linked meta-data were limited, the race and ethnicities of these samples are unknown. While Caucasians are likely to be over-represented in these samples, other ethnicities and racial groups are likely included which may explain the lower-than-expected prevalence of HαT (3% v. 5.6%).

Using the same method, we next examined the expression levels of αWT-, αDUP-, and β-tryptase transcripts in these and other public datasets. This analysis confirmed over-expression of both αWT-, αDUP-tryptase transcripts relative to β-tryptase (Fig. 3a). While over-expression of αWT-tryptase was also observed in our primary gene expression data, it is possible that some δ-tryptase reads could align to the αWT-tryptase consensus sequence, due to haplotype sequence homology, but not to αDUP-tryptase where the linked haplotype is not observed in linkage with δ-tryptase. Therefore, misalignment of δ-tryptase reads could lead to a potential over-estimation of αWT-tryptase expression levels. To exclude this possibility, we examined BST levels – which only measure total secreted α- and β-tryptases – among individuals without HαT. As reported (21), we saw a modest but significant positive correlation between BST levels and increasing αWT-tryptase copy number at TPSABI (Fig. 3b). Thus, consistent with the RNA sequence read alignment results, αWT does appear to be modestly over-expressed relative to β-tryptases. Because of the variable levels of total tryptase transcript in any given sample, we next employed regression analysis to quantify the relative expression of αDUP-tryptase to αWT-tryptase more precisely. Consistent with all our prior data, αDUP-tryptase was found to be over-expressed relative to αWT-tryptase by a factor
of 5.2-fold (95% CI 4.864 to 5.499) in these datasets (R²=0.8, P<0.0001) (Fig. 3c). Thus, in silico, in vitro, and in vivo findings collectively indicate that elevated BST in HαT results from basal over-expression of wild-type α-tryptase at replicated TPSAB1 loci linked to co-inherited over-active promoter elements.

**Modeling BST levels based upon TPSAB1 replication number redefines clinical reference ranges**

Using this knowledge, we created a data-driven model from 204 individuals with normal TPSAB1 copy number and 309 individuals with HαT in whom there was no clinical evidence of clonal mast cell disease (MCD), to generate prediction intervals with upper limits for predicted BST levels based upon TPSAB1 replication number (Fig. 3d) and created on online application for clinical use: Basal Serum Tryptase Clinical cut-off Assigned by Locus Copy number of UTR-Linked element and Associated TPSAB1 Encoded Replication (BST CALCULATOR) available at: [https://bst-calculator.niaid.nih.gov/](https://bst-calculator.niaid.nih.gov/). These thresholds can be used to determine where a bone marrow or further clinical evaluation may be needed for individuals with or without HαT when another clinical indication for such evaluation may be lacking.

This model not only redefined clinically meaningful upper limits for BST levels among individuals with HαT, but it also for the first time defines a clinically actionable upper limit among individuals without HαT as 11.4 ng/mL (Table 3). Remarkably, this has been the upper limit of normal for BST commonly used in most clinical laboratories base. Moreover, based upon the data, this cut-off of 11.4 ng/mL would appear to be valid in defining a clinically abnormal elevated BST when considering the diagnosis of clonal mast cell disease - as opposed to the currently accepted 20 ng/mL - provided that HαT has been ruled out by tryptase genotyping.
This model fit the primary data well ($R^2=0.76$, $P<0.001$) with only one individual being identified in the entire primary dataset as outside the 99.5% threshold (Fig. 3d; Table 3; Fig. S6) and consistent with in vitro and in vivo data, demonstrated linearity on the log-transformed scale (Fig. S6). Interestingly, this individual who did not conform to the model, had provocative clinical phenotype which included chronic spontaneous urticaria and angioedema as well as syncopal episodes with alcohol consumption that had prompted a bone marrow biopsy prior to genotyping. While bone marrow sections demonstrated left-shifted megaloblastoid erythroid hyperplasia, focal areas with increased pronormoblasts, and occasional dyserythropoietic forms, allele-specific PCR for KIT p.D816V was negative, no mast cell aggregates or aberrant markers (i.e., CD2, CD25) were identified, and no definitive diagnosis was made.

Finally to prospectively validate the modeled cut-off clinically, we measured BST levels in an additional 130 individuals with HαT caused by TPSAB1 duplications in whom BST levels were unknown and no other clinical indication for bone marrow biopsy was present. In all 130 prospectively examined individuals, BST levels fell below the 99.5% cut-off (median 15.1 ng/mL, range 8.0-35.0 ng/mL).

**Application of tryptase genotyping and modeled BST levels improves biomarker utility**

To further examine the real-world implications of applying this new model clinically, we analyzed all samples sent or referred to the NIH clinical center for tryptase genotyping from 2016-2021; 377 of these samples had not been used to construct the model. None of the participants had kidney disease or other clinical presentations consistent with genetic disorders known to be associated with elevated BST (13). Of these, 409 samples were identified with BST $>11.4$ ng/mL; 285 of these had HαT, 98 had MCD, 21 had both HαT and MCD, 2 individuals had hypereosinophilic
disorders - one of whom also had HαT - and 4 individuals did not have a clear diagnosis at the time of genotyping. Thus, 98.8% (404/409) of individuals with BST >11.4 ng/mL could be accounted for by HαT and/or MCD (Fig. 3e). Moreover, 97.3% (255/262) of samples with BST levels between 8-11.4 ng/mL were associated with HαT and/or MCD, suggesting that even BST >8 ng/mL is uncommon among individuals without these conditions. Consistent with these data is the modeled 95% threshold among individuals without HαT or MCD of 7.99 ng/mL, indicating that less than 5% of the total population would be predicted to have BST >8 ng/mL.

The 5 individuals with BST >11.4 ng/mL who did not have HαT or evidence of MCD underwent paired peripheral blood exome and deep exome sequencing (300x) of granulocyte and mononuclear bone marrow aspirate fractions. In all five individuals, variants in genes commonly mutated in hematological malignancies (e.g., TET2, RUNXI, CEBPA, MAP2K1, NOTCH2, KMT2C) were identified (Table S6); in one individual this was associated with idiopathic hypereosinophilic syndrome. Together with those individuals meeting criteria for MCD, these data demonstrate that BST >11.4 ng/mL in the absence of HαT frequently identifies indolent clonal mast cell disease - 95% (98/103) of our referral cohort - or more rarely identifies patients with genetic variants suggestive of occult or evolving myeloid neoplasms.

Of note, 38.2% (55/144) samples with BST >20ng/mL were a result of HαT alone. Of those 55 participants, 54 were less than their respective threshold based on replication number. If tryptase genotyping was not considered, and instead the current minor criterion for SM of BST >20 ng/mL was used as a sole indication for bone marrow evaluation, 54 of these individuals could have been subjected to unnecessary invasive bone marrow examination. This is particularly relevant among individuals with two or more TPSAB1 replications, where median BST levels were
>20 ng/mL and can be >100 ng/mL, as seen in two related individuals with ten *TPSAB1*
replications in cis (Fig. 2b, 3d; Table 3).

Finally, of the 21 individuals identified with both HαT and MCD (Table 4), 10 individuals had indolent SM (ISM), 8 of whom had BST levels above the 99.5% upper prediction limit dictated by *TPSAB1* replication number. Ten individuals had BST levels below the 99.5% upper prediction limit based upon *TPSAB1* replication number in our model; 6 with monoclonal mast cell activation syndrome (MMAS), 2 with ISM, 1 with diffuse cutaneous mastocytosis, and 1 with well-differentiated ISM. All these individuals had clinical implications for a bone marrow beyond the BST level and would have received a bone marrow regardless. Only one of these individuals - with the rare phenotype of well-differentiated ISM - would no longer have met clinical criteria for their diagnosis if the upper prediction limit was used as a minor criterion for the diagnosis of SM (rather than the current level of >20 ng/mL). Together, these data demonstrate that where tryptase genotyping is clinically available, these new reference ranges based upon *TPSAB1* replication number can be applied to the work-up of patients with elevated BST, establishing robust thresholds for which patients should undergo more extensive work-up including bone marrow aspiration and biopsy regardless of clinical presentation or symptomatology, and as an individualized minor criterion for the clinical diagnosis of SM.

**DISCUSSION**

Based upon our findings, HαT is a genetic trait best described as a naturally occurring over-expression model, whereby a conserved over-active protomer is co-inherited with α-tryptase-encoding *TPSAB1* gene replications driving basal over-expression of α<sup>DUP</sup>-tryptase that when translated, is indistinguishable from other α-tryptase proteins (i.e., α<sup>WT</sup>-tryptase). Using this
knowledge and clinical data, we have generated a model that redefines clinically meaningful cut-offs for the upper limit of normal for serum tryptase levels based upon tryptase genotype. Moreover, we have demonstrated the clinical utility of using these individualized values in the evaluation of patients with elevated serum tryptase levels and in the diagnosis of systemic mastocytosis.

We acknowledge that MCD is not the only clonal myeloid neoplasm associated with elevated BST, rather it is the predominant one referred to our centers. Moreover, several studies have now demonstrated an increased prevalence of HaT among symptomatic individuals with SM (6-8) - likely due to potentiation of mast cell-mediated symptoms and reactions. Additional prospective studies will be helpful in validating the use of these new reference values in additional patients with HaT, MCD and other myeloid neoplasms. Severe renal dysfunction can also affect BST levels, as can other parasitic diseases which are rare in most Western countries (15). These are also not represented or evaluated in this study and should be considered when applying this model to patients. However, the model was generated based upon individuals without clinical signs or symptoms of these disorders. Thus, we expect these new reference values to remain robust in their predictive value, as has been suggested by a recent retrospective study of a regional health system where 93% (54/58) of individuals with BST >11.4 ng/mL had HaT, a myeloid neoplasm, or chronic kidney disease (22).

The data used to generate our model is largely from individuals with two or fewer replications. Thus, while available data for higher order copy number are consistent with our model, and indeed were incorporated into it, the generated cut-offs at high TPSAB1 replication number (e.g., ≥3) are largely extrapolations. The a priori evidence demonstrated a linear relationship between BST and replication number providing support for this extrapolation.
However, there is less certainty about the upper limits of BST levels past two TPSAB1 replications, and cut-offs for individuals with higher order copy number may require refinement as additional data become available.

To our knowledge, the application of genotypic information to determine clinical reference ranges in a personalized manner has not previously been used in laboratory medicine. However, many common clinical laboratory measures (e.g., immunoglobulin levels) are not normally distributed, and in some cases outliers may be similarly determined by heritable traits. Thus, we anticipate this kind of precision approach to clinical laboratory medicine will expand in the future as next generation sequencing becomes increasingly utilized in standard clinical practice.

**METHODS**

**Study participants and samples**

Patients, family members, and healthy volunteers provided informed consent on IRB-approved research protocols led by investigators with expertise in Allergy/Immunology at institutions that specialize in clonal and non-clonal mast cell disorders designed to study mastocytosis, or genetic diseases affecting the immune system at the NIH Clinical Center (NCT00852943, NCT01164241, NCT00044122, NCT00001756, NCT00719719), Antwerp University Hospital, Belgium (B300201525454), Verona University Hospital, Italy (protocol No. 39620), the University of Florida (IRB 201702274), the University of Mississippi Medical Center (IRB 2019-0082), or University Clinic Golnik, Slovenia (KME 150/09/13). All study participants had BST levels measured and tryptase genotyping performed (N=1,178). Among participants referred for evaluation or diagnosed with a mast cell-associated disorder (N=575), complete history and physical examinations were performed.
Bone marrow biopsy and aspirate

Individuals presenting with signs or symptoms suggestive of a clonal mast cell disease (13) had additional clinical work-up that included bone marrow biopsy and aspirate. Immunohistochemistry of bone marrow sections was performed for enumeration and characterization of mast cells (KIT and tryptase), and evaluation of CD2 and CD25 expression in aspirate and/or tissue section. Allele-specific PCR for KIT p.D816V in peripheral blood and bone marrow was also performed in these patients. In the rare case that an individual declined bone marrow biopsy, peripheral blood KIT p.D816V was performed by allele-specific or ddPCR. If the result was positive, these individuals were considered to have clonal mast cell disease and included. Whereas, if KIT p.D816V screening of peripheral blood was negative in an individual with an incomplete work-up for a suspected myeloid neoplasm, such individuals were excluded from the study.

Total basal serum tryptase quantification

Total basal serum tryptase (BST) levels were measured using the commercially available ImmunCAP assay (Thermo Fisher Scientific, Phadia AB, Uppsala, Sweden) or ELISA as described (16), performed in CLIA-certified laboratories (Mayo Clinic, Rochester, NY and Virginia Commonwealth University, Richmond, VA).

Bacterial artificial chromosome (BAC) library generation

Genomic DNA was isolated from PBMC of an individual homozygous for ααβ tryptase alleles. HMW DNA fragments were then partially digested with the restriction enzyme BamHI and size selected prior to ligation of fragments into the pCC1BAC vector (Epicentre Biotechnologies, Inc., Madison, WI, for use under a CC0 license. This article is a US Government work. It is not subject to copyright under 17 USC 105 and is also made available for use under a CC0 license.)
Madison, WI) and transformation of DH10B E. coli cells (17, 18). The clones were robotically chosen and replicated. Replicated clone copies were used as a source plate for constructing nylon filters. The filter hybridization was carried out using the DIG method for filter hybridization as described (19) (Amplicon Express, Pullman, WA). Verification of putative clones was carried out by conventional and digital droplet PCR (ddPCR).

BAC sequencing and assembly

Genomic DNA from selected clones was isolated (20) and sequenced using Single-molecule real-time sequencing (SMRT) sequencing technology (PacBio, Menlo Park, CA) at DNA Link (South Korea). De novo assembly was accomplished using HGAP2 and the Canu assembler for PacBio on SMRT Analysis platform version 2.2.0 and the SMRTPortal: SMRTAnalysis build 133377, Daemon version v2.2.0 build 132105, SMRTpipe version v2.2.0 build 132739, SMRT Portal version v2.2.0 build 133335, SMRT View version v2.2.0 build 132578.

Bioinformatic analyses

Paired-read tryptase haplotype analysis

Genome sequence reads mapping to the tryptase locus were re-aligned to a 142-bp region (Table S1). Detailed alignments for each read were parsed to extract the sequence at each of the five specified variable positions if covered by the read. Haplotype assignments were determined based on the collection of positional sequence assignments determined for each read pair. Generated haplotypes were then correlated to tryptase genotype and linkage to inherited alleles was confirmed when pedigrees were available.
**RNA-seq dataset analyses**

Using a 39-bp consensus sequence (Table S2), gene expression datasets (see Accession codes for list of datasets) were reanalyzed using NCBI Magic-BLAST (https://ncbi.github.io/magicblast/), in order to map reads to 3 gene isoforms: \( \alpha^{\text{DUP}} \), \( \alpha^{\text{WT}} \), and \( \beta \)-tryptase. Samples were designated as tryptase-positive, tryptase-negative, \( \alpha^{\text{DUP}} \), \( \alpha^{\text{WT}} \), and \( \beta \)-tryptase only based on the criteria of having qualified reads that mapped exactly without any mismatches to the interval between 25 bp-45 bp from 5’ end of the transcripts (see Supplemental Methods).

**Promoter amplification and cloning**

Promoter amplicons were generated from genomic DNA from individuals with increased \( TPSAB1 \) copy number using primers designed on conserved sequences present in all identified alpha- and beta-tryptase sequences (FWD – GGGCAAGTCCACAGGGAGCT; REV – CTGGGGAGCAAGGGAGGAGCA) in order to amplify all sequences between the ATG start site and a conserved region ~1-2 kb 5’ of the variably expanded repeat region (Fig. 1). Amplification was confirmed by gel electrophoresis and clones of products were generated using the TOPO® Cloning Kit (Thermo, Waltham, MA) and transformed into One Shot® TOP10 Chemically Competent E. coli (Thermo). Single colonies were selected, confirmed to contain single intact clones by PCR and gel electrophoresis, and Sanger sequenced using the FWD and REV, as well as two additional internal primers (TGCAGGTGCAACCCCAGGA and TCCTGGGGTTGCACTGCA).

For \( \alpha^{\text{DUP}} \) and \( \alpha^{\text{WT}} \)-tryptase-specific promoter cloning, the reverse primer (GACGATACCCGCTTGCTGCAG) was located across alpha-tryptase-specific exonic sequence used for the isoform-specific genotyping assay, and for \( \alpha^{\text{WT}} \)-tryptase the universal FWD primer
was used and for $\alpha^{DUP}$-tryptase a different sequence normally only seen linked to beta-tryptase was used (GGGCAAGTCCACACGGGAGCT). Resulting amplicon size was confirmed by gel electrophoresis (Fig. S2).

**Reporter assay**

Sequence-validated clones corresponding to $\alpha^{DUP}$- and $\alpha^{WT}$-tryptase promoters identified in the BAC assembly were subcloned into the reporter plasmid pDD-AmCyan1 (Takara Bio USA, Inc., Mountain View, CA) and Sanger sequence verified. MonoMac-6 cells were transfected with the reporter plasmid pDD-AmCyan1 containing $\alpha^{DUP}$- or $\alpha^{WT}$-tryptase promoter clones by electroporation using Cell Line Nucleofector™ Kit V (Lonza, Basel, Switzerland) using the setting U-005 according to manufacturer instructions. Cells were cultured in standard media in the presence of Shield1 (Takara Bio) according to manufacturer instructions and basal fluorescence was measured at indicated time points and recorded using an LSR Fortessa (BD Biosciences) and analyzed using FlowJo (Treestar, Ashland, OR).

**Basophil isolation**

Following isolation of PBMC via density centrifugation, basophils were negatively selected using the MACS® Diamond Basophil Isolation Kit (Miltenyi Biotec, San Diego, CA).

**Cell culture**

**Primary mast cells**

Peripheral blood mononuclear cells (PBMCs) were isolated using density centrifugation and CD34$^+$ cells were positively selected using the ferromagnetic bead-based MACS® system.
(Miltenyi) and the number of cells was quantified. Mast cells were then differentiated under the established conditions described (21).

**Cell lines**

MonoMac-6 cells were cultured in RPMI 1640 + 10% FBS + 2 mM L-glutamine + non-essential amino acids + 1 mM sodium pyruvate + 10 µg/ml human insulin).

**Droplet digital PCR (ddPCR)**

**Tryptase genotyping**

Alpha- and beta-tryptase sequences at *TPSAB1* and *TPSB2* were genotyped as described (5) using droplet digital PCR (ddPCR). Briefly, genomic DNA was extracted from PBMCs, cell lines, or obtained from the HapMap biorepository and restriction endonuclease treated. Custom primer/probe sets specific for alpha- and beta-tryptase were then employed using the PrimePCR ddPCR copy number reference *AP3B1*, according to the manufacturer's specifications (Bio-Rad; Hercules, CA).

**Gene expression**

Total RNAs were extracted using Trizol reagent (Thermo) and the RNEasy Mini Kit (Qiagen) from basophils, primary cultured mast cells, or cell lines under stated conditions and reverse-transcribed using SuperScript III First-Strand Synthesis System (Thermo). ddPCR was performed on a QX200 system (Bio-Rad) according to the manufacturer’s instructions using custom primer/probe sets for tryptase isoforms as indicated (Table S3) to quantify gene expression,
normalizing to *HPRT1* and *TBP* expression quantified using commercially available ddPCR primer/probe sets (Bio-Rad).

**Promoter detection assay**

A custom primer/probe set was designed to amplify the proximal promoter linked to alpha-tryptase (Fig. S4). The forward primer corresponds to a conserved sequence present in all identified α- and β-tryptases. The reverse primer hybridizes to sequences only present in alpha-tryptases. The probes compete for hybridization with the C>T variant that defines α<sup>DUP</sup>- and α<sup>WT</sup>-tryptase-specific sequences, respectively. To successfully amplify the large 71% GC-rich amplicon, the reaction was run with 10% 1M Betaine (Sigma, Saint Louis, MO) under the following conditions: 95 °C 10 min, 96 °C 30 s, 66 °C 1 min (Ramp 1.5 °C), 50 cycles, 96°C 10 min.

**Exome sequencing and bioinformatic analysis for clonal variants**

Bone marrow aspirates were fractionated using density centrifugation into granulocyte and mononuclear fractions using a double gradient 1.077 g/ml (Histopaque-1077) and 1.119 g/ml (Histopaque-1119) (Sigma, St. Louis, MO). Genomic DNA was extracted and libraries were prepared with the Twist Biosciences Comprehensive Exome capture kit. Deep (>300x) exome sequencing was subsequently performed using the Illumina HiSeq 2500, on both the granulocyte and mononuclear bone marrow fractions. Paired peripheral blood samples from the same individuals were also exome sequenced consistent with clinically recommended standards (22).

Raw fastq files were trimmed for quality and adapter contamination using Trimmomatic v0.39 (23) and mapped to the human hg38 reference genome using BWA-MEM v0.7.17 (http://bio-bwa.sourceforge.net/). PCR duplicates were marked using Samlaster v0.1.2.5 (24), and
GATK v4.1.9.0 was used to perform base recalibration. For the peripheral blood samples, germline variation was called using GATK v4.1.9.0 and following the recommended Best Practices (https://gatk.broadinstitute.org/hc/en-us/articles/360035535932-Germline-short-variant-discovery-SNPs-Indels-). For granulocyte and mononuclear fractions, somatic variants were called using 2 approaches. First, paired somatic calling with the PBMC sample treated as germline was performed using 3 somatic detection tools: 1) MuTect2 from GATK v4.1.9.0 and following the recommended Best Practices (https://gatk.broadinstitute.org/hc/en-us/articles/360035531132--How-to-Call-somatic-mutations-using-GATK4-Mutect2), 2) Strelka v2.9.0 (25), and 3) mutect v1.1.7 (26). In addition to paired calling, we also performed tumor-only calling with a combination of MuTect2 GATK v4.1.9.0, mutect v1.1.7, and Vardict v1.7.0 (27). The latter approach of performing somatic variant detection without including the paired PBMC sample was done to account for the fact that somatic variants in the bone marrow granulocytes and bone marrow mononuclear cells could also be present at low fractions in the PBMC sample and therefore get excluded as contaminating germline variants in the paired calling. For both approaches, variants were then merged across all callers and annotated with VEP v104 (28) for downstream analysis. Reported variants were prioritized based upon commonly implicated genes in the development of myeloid neoplasms.

Statistical modeling and analyses

To predict BST levels based upon tryptase genotypes - specifically determined by TPSAB1 replication number (including those without replications) - a log transformation was applied to the BST levels because the values were not normally distributed. Additionally, we log-transformed the replication number, because transforming BST back to the normal scale resulted in a model fit
that was approximately linear, with a near constant increase in BST for each increase in replication number; this supported results from early clinical data and data generated from in silico and in vitro experiments. We also added 0.5 to replication number prior to log transformation in order to include individuals without $TPSAB1$ replications in the model, thus allowing us to also determine a cutoff for individuals without $\alpha$T. Thus, we used a linear regression model to predict \log(BST) from \log(replication number plus 0.5), and from that, created an upper one-sided 99.5% prediction interval. The 99.5% threshold was chosen as it provides high specificity where very few individuals will be identified as false positives and correlated well with the predicted prevalence of elevated BST in the absence of $\alpha$T based upon available population data (10). The model was developed using R version 3.6.3.

For clinical and experimental data, Mann-Whitney, Kruskal-Wallis, and paired and unpaired two-tailed t-tests, were used where appropriate to test significance of differences, prevalence, or deviation using Prism (GraphPad).

**Basal Serum Tryptase (BST) calculator code**

We developed an online calculator the Basal Serum Tryptase Clinical cut-off Assigned by Locus Copy number of UTR-Linked element and Associated $TPSAB1$ Encoded Replication (BST CALCULATOR) with Shiny R framework: [https://bst-calculator.niaid.nih.gov/](https://bst-calculator.niaid.nih.gov/). The code is available at: [https://github.com/niaid/BST-calculator](https://github.com/niaid/BST-calculator)

**URLs**


Plink, http://pngu.mgh.harvard.edu/purcell/plink/

Genome Analysis Toolkit (GATK), http://www.broadinstitute.org/gatk/

SAMtools, http://www.htslib.org/

Accession Codes

Sequencing data were obtained from the National Center for Biotechnology Information Sequence Read Archive for the following BioProjects: hereditary alpha-tryptasemia: PRJNA342304; individuals not selected by tryptase genotype: PRJNA208369 PRJNA219425 PRJNA232669

PRJNA252605 PRJNA253059 PRJNA254943 PRJNA257389 PRJNA258216
PRJNA261011 PRJNA261251 PRJNA263242 PRJNA266512 PRJNA266572
PRJNA270371 PRJNA271942 PRJNA274028 PRJNA274360 PRJNA275801
PRJNA279249 PRJNA280990 PRJNA283839 PRJNA289905 PRJNA291619
PRJNA292690 PRJNA293555 PRJNA296379 PRJNA301173 PRJNA301364
PRJNA310988 PRJNA315611 PRJNA317535 PRJNA318253 PRJNA319220
PRJNA326113 PRJNA327986 PRJNA330840 PRJNA340161 PRJNA342177
PRJNA343985 PRJNA354367 PRJNA358081 PRJNA369563 PRJNA369684
PRJNA373887 PRJNA376200 PRJNA377555 PRJNA378385 PRJNA384963
PRJNA388978 PRJNA389466 PRJNA392116 PRJNA395367 PRJNA395589
PRJNA399103 PRJNA400331 PRJNA407731 PRJNA415746 PRJNA428940
PRJNA214592 PRJNA229548 PRJNA278364 PRJNA278767 PRJNA263397

Referenced GenBank Accession Codes
GRCh37/hg19 NCBI assembly: GCF_000001405.13; tryptase locus assemblies: AC226137.3

AC120498.2 CTD-2503P16 CHM13 AL031704.24 AE006466.1 CHM1 AC240106.3

AC238650.2 AC213746.1 GM24385 NA12878 AF098328.1

Acknowledgements: The investigators thank the patients, their families, and healthy volunteers who contributed to this research. We also thank Adam Phillippy for sharing his expertise on complex genetic structural rearrangement.
References


Figure 1. TPSAB1 duplications occur at 16p13.3 and are linked to an expanded promoter associated with αDUP-tryptase over-expression. (a) Location and orientation of the 15-kB tandem duplication of TPSAB1. (b) Alignment of duplicated alpha- (αDUP), wild-type alpha (αWT), and beta1-tryptases with unique 5’-variants and size of promoter repeat regions indicated; repeat motif within promoter regions (c). Relative total α- to β-tryptase gene expression (d) and αDUP-relative to αWT-tryptase gene expression (e) in ex-vivo basophils and cultured primary mast cells (right). *P<0.05; **P<0.005.
Figure 2. The expanded promoter at duplicated TPSAB1 is conserved and has increased basal activity. (a) Normalized fluorescent-reporter activity in unstimulated MonoMac-6 cells transfected with promoters cloned from endogenous ($\alpha^{WT}$) or replicated ($\alpha^{DUP}$) TPSAB1 loci. (b) Basal serum tryptase (BST) from individuals grouped by TPSAB1 replication number. Relative allelic frequency (c) and ratios (d) of $\alpha^{WT}$- (y-axis) and $\alpha^{DUP}$-tryptase (x-axis) associated 5'-variants determined by droplet digital PCR; dashed lines indicate predicted $\alpha^{WT}$:$\alpha^{DUP}$ ratios by genotype. *P<0.05
Figure 3. Modeling serum tryptase levels based upon genotype improves clinical utility. (a) Normalized (left) and total (right) read counts for reads aligning exactly to the 39-bp consensus sequences that identify β-, αWT- and αDUP-tryptase. (b) BST levels among individuals with conserved 4n tryptase copy number (combined from TPSAB1 and TPSB2). (c) Regression analysis of relative expression levels of αDUP (y-axis) and αWT (x-axis) tryptase transcripts. (d) Prediction intervals (PI) for BST levels based upon TPSAB1 replication number. (e) Prevalence of HaT, clonal mast cell disease (MCD), and those without either, among individuals referred with BST levels above the predicted upper limit of normal (>11.4 ng/mL). *P<0.01; **P<0.005; ***P<0.0001
Table 1. Assemblies of the human tryptase locus at 16p13.3.

<table>
<thead>
<tr>
<th>TPSB2 haplotype orientation</th>
<th>TPSAB1&lt;sub&gt;REP&lt;/sub&gt; isofrom</th>
<th>TPSAB1 haplotype orientation</th>
<th>TPSD1 haplotype orientation</th>
<th>Corresponding Assembly</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACCT REV beta3</td>
<td>-</td>
<td>-</td>
<td>GACCC FWD beta1</td>
<td>GRCh37/hg19</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GACCC REV beta1</td>
<td>AC226137.3</td>
</tr>
<tr>
<td>CACCT REV beta3</td>
<td>-</td>
<td>-</td>
<td>GACCC FWD beta1</td>
<td>AC120498.2</td>
</tr>
<tr>
<td>CACCT REV beta3</td>
<td>-</td>
<td>-</td>
<td>CACCT REV beta3</td>
<td>CTD-2503P16</td>
</tr>
<tr>
<td>GACCT REV beta2</td>
<td>-</td>
<td>-</td>
<td>GGTTC FWD alpha</td>
<td>AE006466.1</td>
</tr>
<tr>
<td>GACCT REV beta2</td>
<td>-</td>
<td>-</td>
<td>GGTTC REV alpha</td>
<td>AC240106.3</td>
</tr>
<tr>
<td>GACCT REV beta2</td>
<td>-</td>
<td>-</td>
<td>GGTTC REV alpha</td>
<td>AC238650.2</td>
</tr>
<tr>
<td>GACCC REV beta1</td>
<td>CACCT REV alpha</td>
<td>GGTTC REV alpha</td>
<td>GGTTC FWD beta1</td>
<td>HαT BAC</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>GGTTC - alpha</td>
<td>AF098328.1</td>
</tr>
</tbody>
</table>

Bold – assembly of the locus from an individual with hereditary alpha-tryptasemia (HαT)
Table 2. Haplotypes identified in linkage with allelic tryptase genotypes.

<table>
<thead>
<tr>
<th>TPSB2 haplotype</th>
<th>isoform</th>
<th>TPSA1&lt;sup&gt;DUP&lt;/sup&gt; haplotype</th>
<th>isoform</th>
<th>TPSA1 haplotype</th>
<th>isoform</th>
<th>TPSD1 haplotype</th>
<th>Allele count</th>
<th>Allelic genotype</th>
<th>Haplotype Frequency (by genotype)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACCT beta3</td>
<td>-</td>
<td>GACCC beta1</td>
<td>GGTTC</td>
<td>172</td>
<td>β/β</td>
<td>52%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CACCT beta3</td>
<td>-</td>
<td>GACCC beta1</td>
<td>CGTTC</td>
<td>152</td>
<td>β/β</td>
<td>46%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCT beta2</td>
<td>-</td>
<td>GACCC beta1</td>
<td>GGTTC</td>
<td>7</td>
<td>β/β</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCT beta2</td>
<td>-</td>
<td>GGTTC alpha</td>
<td>GGTTC</td>
<td>65</td>
<td>α/β</td>
<td>96%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCC beta1</td>
<td>CACCT</td>
<td>GGTTC alpha</td>
<td>CGTTC</td>
<td>3</td>
<td>α/β</td>
<td>4%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCC beta1</td>
<td>CACCT</td>
<td>-</td>
<td>CGTTC</td>
<td>15</td>
<td>αα/β</td>
<td>63%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCC beta1</td>
<td>CACCT</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>αα/β</td>
<td>21%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCC beta1</td>
<td>CACCT</td>
<td>GGT (6delC)C* alpha</td>
<td>GGTTC</td>
<td>2</td>
<td>αα/β</td>
<td>8%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GACCC beta1</td>
<td>CACCT</td>
<td>GGTTC alpha</td>
<td>GGTTC</td>
<td>2</td>
<td>αα/β</td>
<td>8%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A single cytosine (C) base deletion was identified 6 bases prior (-6) to the terminal C of the haplotype 747748*
Table 3. Measured and predicted BST levels based upon *TPSAB1* replication number encoding α-trypstase.

<table>
<thead>
<tr>
<th>TPSAB1 replication number</th>
<th>N</th>
<th>Median BST</th>
<th>SD BST</th>
<th>Modeled reference values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fit</td>
</tr>
<tr>
<td>0</td>
<td>204</td>
<td>4.15</td>
<td>2.12</td>
<td>4.2831</td>
</tr>
<tr>
<td>1</td>
<td>247</td>
<td>13.60</td>
<td>5.55</td>
<td>13.6363</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>22.45</td>
<td>6.40</td>
<td>23.3642</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>26.00</td>
<td>NA</td>
<td>33.3110</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>40.50</td>
<td>12.61</td>
<td>43.4148</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>53.6420</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>87.00</td>
<td>NA</td>
<td>63.9708</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>74.3862</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>84.8774</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>95.4357</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>133.00</td>
<td>32.53</td>
<td>106.0544</td>
</tr>
</tbody>
</table>

BST - basal serum tryptase; SD - standard deviation; PI - prediction interval
Table 4. Individuals with MCD and HgT.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Age</th>
<th>Gender</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>MC aggregates*</th>
<th>Spindle†</th>
<th>KIT p.D816V‡</th>
<th>CD2 or CD25‡</th>
<th>BST &gt;20ng/mL§</th>
<th>Above 99.5% PI</th>
<th>New Diagnosis†</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISM</td>
<td>66-70</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>236</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>20-25</td>
<td>M</td>
<td>4</td>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>107</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>66-70</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>88.7</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>66-70</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>yes</td>
<td>negative</td>
<td>positive</td>
<td>70</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>10-15</td>
<td>F</td>
<td>2</td>
<td>3</td>
<td>yes</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>60.1</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MMAS</td>
<td>70-75</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>no</td>
<td>negative</td>
<td>positive</td>
<td>52.2</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>50-55</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>51</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>70-75</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>positive</td>
<td>negative</td>
<td>42</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MMAS</td>
<td>66-70</td>
<td>F</td>
<td>2</td>
<td>3</td>
<td>no</td>
<td>no</td>
<td>negative</td>
<td>positive</td>
<td>38.8</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM (WD)</td>
<td>30-35</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>no</td>
<td>negative</td>
<td>negative</td>
<td>37.9</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>ISM</td>
<td>10-15</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>37.6</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>MMAS</td>
<td>20-25</td>
<td>M</td>
<td>6</td>
<td>2</td>
<td>no</td>
<td>no</td>
<td>positive</td>
<td>negative</td>
<td>37</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>DCM</td>
<td>6-10</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>NA</td>
<td>NA</td>
<td>negative</td>
<td>positive</td>
<td>32.6</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>ISM</td>
<td>56-60</td>
<td>F</td>
<td>2</td>
<td>3</td>
<td>no</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>32.2</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>MMAS</td>
<td>50-55</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>negative</td>
<td>negative</td>
<td>27.3</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>MMAS</td>
<td>60-65</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>no</td>
<td>positive</td>
<td>negative</td>
<td>26.8</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>ISM (WD)</td>
<td>26-30</td>
<td>F</td>
<td>2</td>
<td>3</td>
<td>yes</td>
<td>no</td>
<td>negative</td>
<td>negative</td>
<td>26.8</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>ISM</td>
<td>30-35</td>
<td>F</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>yes</td>
<td>positive</td>
<td>positive</td>
<td>23.7</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>MMAS</td>
<td>60-65</td>
<td>M</td>
<td>2</td>
<td>3</td>
<td>no</td>
<td>no</td>
<td>positive</td>
<td>negative</td>
<td>23.1</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>MMAS</td>
<td>56-60</td>
<td>M</td>
<td>3</td>
<td>2</td>
<td>no</td>
<td>no</td>
<td>positive</td>
<td>negative</td>
<td>20.3</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>MMAS</td>
<td>66-70</td>
<td>F</td>
<td>2</td>
<td>3</td>
<td>no</td>
<td>no</td>
<td>positive</td>
<td>negative</td>
<td>18.8</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

ISM - indolent systemic mastocytosis; WD - well-differentiated; MMAS - monoclonal mast cell activation syndrome; DCM - diffuse cutaneous mastocytosis; *major and †minor criteria for the diagnosis of SM: §a new diagnosis being assigned based upon application of the 99.5% threshold for BST, rather than >20ng/mL; §in peripheral blood only. The individual with a change in diagnosis based upon genotype-defined BST reference values is indicated in bold italic font.