Published on Apr 23, 2025 by Arcadia Science **DOI**: 10.57844/arcadia-77d4-1c5d

How confident should we be in potential targets of tick protease inhibitors predicted by AlphaFold-Multimer?

We want to predict the targets of tick effectors to identify new therapeutic targets for skin diseases. We ran a case study using AlphaFold-Multimer to predict the targets of tick protease inhibitors, but we aren't sold on our method. What other approaches should we consider?

Contributors (A-Z)

Audrey Bell, Adair L. Borges, Feridun Mert Celebi, Megan L. Hochstrasser, Elizabeth A. McDaniel, Austin H. Patton, Taylor Reiter, Emily C.P. Weiss

Version 1 · May 02, 2025

Purpose

Human parasites have evolved to be expert manipulators of our biology, and we're interested in the potential for ectoparasites like ticks to point us to the high-leverage therapeutic targets for dermatological disease. Computational protein–protein

interaction prediction methods have the exciting potential to allow us to map the targets of tick effectors at an unprecedented scale.

However, when we conducted a small case study mapping the targets of a family of tick protease inhibitors, we weren't sure how to interpret our results. While some of our hits seemed logical (i.e., present in the skin and connected to itch and inflammation), we also got many hits to non-physiologically relevant proteases, such as human digestive enzymes. We're sharing this case study to get feedback from others who have used similar tools or asked related questions.

- This pub is part of the **project**, "Ticks as treasure troves: Molecular discovery in new organisms. Visit the project narrative for more background and context.
- **Data**, including protein sequences, annotations, structures, our full ProteinCartography analysis, and our raw AF-Multimer results, are on Zenodo.
- **Additional data**, including phylogenetic profiling results, D-SCRIPT results, and AF-Multimer summary results, plus all associated **code**, can be found in our GitHub repo.

We've put this effort on ice!

#TechnicalGap

We identified host proteases predicted to be targeted by tick protease inhibitors, but some targets seem more biologically plausible than others. Since we don't have a method to evaluate our false-positive rate, we aren't sure how confident we should be in our results. Without understanding this, we can't justify moving the project forward.

Learn more about the Icebox and the different reasons we ice projects.

What are we trying to do?

Ticks are professional manipulators of skin. When ticks take a blood meal, they inject pharmacologically active molecules into host skin to precisely control complex host processes such as coagulation, inflammation, wound healing, itch, and pain **[1]**. Ticks also feed for days to weeks **[2]**, requiring prolonged and stable control of these processes. We think this is translationally useful, as many skin diseases are characterized by chronic inflammation, itch, and pain. There's substantial unmet need in chronic skin conditions, in part driven by the complexity of the underlying biology of these processes **[3]**. We see this as a target selection problem: it's not clear which human proteins or pathways represent the most tractable point to manipulate therapeutically to obtain a durable patient response.

We reasoned that ticks might be able to point us to the highest-leverage therapeutic targets for controlling these processes. However, there are hundreds to thousands of putative tick effectors and an even larger number of potential targets in the human proteome. To this end, we have begun exploring computational approaches to predict cross-species protein–protein interactions, with the dream of comprehensively predicting the targets of tick salivary effectors across the human proteome. We decided to start small, with a case study predicting the targets of a family of tick protease inhibitors. It's well known that protease inhibitors represent a large fraction of tick salivary effectors **[4]**, but most tick protease inhibitors haven't been matched to their targets. Proteases are a large class of enzymes involved in many processes that play evolutionarily significant roles in the feeding ecology of ticks (e.g., blood coagulation **[5]**, wound healing **[6]**, inflammation and immunity **[7]**, itch **[8]**, pain **[9]**, etc). In this case study, we used protein–protein interaction (PPI) prediction to identify the targets of a single family of protease inhibitors.

We first selected a family of tick trypsin-inhibitor-like (TIL) protease inhibitors to use in our case study. Then, we tested out PPI prediction tools D-SCRIPT **[10]** and AlphaFold-Multimer (AF-Multimer) **[11]** to identify targets of these inhibitors. Last, we tried to contextualize our results to evaluate whether these tools gave us actionable insights. This is where we got stuck: while we can predict targets, we have no idea how reliable we should consider these hits. We're sharing this case study to solicit opinions and feedback from others who have worked on similar problems.

Our approach so far

As our goal was to identify the targets of a single, promising protease inhibitor gene family, we first identified protease inhibitor genes and then connected them to tick gene families that we'd previously identified **[12]**. We then used phylogenomic trait association tests **[12]** to rank protease inhibitor families according to the strength of their association with host detection suppression by the parasites. Using the most strongly positively associated protease inhibitor family, we tried out protein–protein interaction prediction tools D-SCRIPT and AF-Multimer to predict the human proteases targeted by a single family of tick protease inhibitors.

Keep reading for more details or skip straight to "The results."

Identifying putative tick protease inhibitors and secreted proteins of unknown function (PUFs) with HMMs

The first step in our analyses was to identify putative tick protease inhibitors in ticks using HMM annotations. We used 15 tick proteomes that we'd downloaded and annotated with DeepSig (v1.2.5) **[13]**, KofamScan (v1.3.0) **[14]**, and eggNOG-mapper (v2.1.10) **[15]** as part of a larger effort to explore the genetic basis of host detection suppression in ticks and other chelicerates **[12]**. We used these annotations to select proteins with annotations that relate to protease inhibitor function. Because some host-interacting protease inhibitors may have been too divergent to receive an annotation from eggNOG-mapper or KofamScan, we also chose to move all secreted proteins of unknown function (PUFs) forward to structural analysis. We considered a protein to be of unknown function if it didn't have any eggNOG-mapper annotation and if it didn't have a KofamScan hit above the KofamScan-recommended threshold. We categorized proteins as secreted if they had the DeepSig feature "Signal peptide." We then removed all proteins > 1,200 amino acids, due to the challenges associated with folding larger proteins.

Check out "Notebook 01" in our GitHub repo for our **protein selection code**.

Structural prediction with ESMFold

At the time of this analysis, the majority of our tick-secreted PUFs and protease inhibitors had no predicted structures. So, we used ESMFold **[16]** to predict the structures for these proteins to use in downstream structural comparisons. To do this, we used a pre-trained EsmForProteinFolding model from the Hugging Face transformers package (v4.45.0) to predict protein structures with our Arcadia Science-hosted ESMFold API. We used the pre-trained model "facebook/esmfold_v1," and we replaced ambiguous residues (marked as "X" in the sequence) with alanine prior to folding. Ambiguous residues arise from errors in sequencing, where bases are left as "N," and translated as X in the amino acid sequence. We chose to replace these unknown amino acids with alanine, a chemically neutral amino acid, to enable folding by ESMFold. The "enque.py" script cleans the sequences and submits the job, whereas the "download_pdbs.py" script fetches the folded results from the ESMFold API. We didn't perform any additional refinement on the predicted structures.

> Access our **predicted structures** in "pi-disco_esmfold_structures.zip" on Zenodo (10.5281/zenodo.15186244).

Downloading UniProt toxin database

To identify novel or divergent protease inhibitor families in ticks, we compared our predicted tick protein structures to the UniProt animal toxin annotation project database of 7,065 manually reviewed toxins. Many of these toxins are protease inhibitors from other venomous species. We downloaded the AlphaFold-generated PDB files of venom toxin proteins from UniProt using "fetch_accession.py" from ProteinCartography **[17]** and downloaded the metadata file using "fetch_uniprot_metadata.py." We were successful in downloading 7,008 structures.

> View the metadata file and list of downloaded structures.

Using ProteinCartography to identify tick protease inhibitors

The next step in our protease inhibitor discovery workflow was to use structural clustering to identify major structural families of protease inhibitors in ticks. We clustered our tick protease inhibitors and secreted proteins of unknown function with a UniProt structural database of venom toxins using ProteinCartography (v0.4.0-alpha) **[17]** in cluster mode using default parameters. This produces Leiden clusters of structurally related proteins. Many of these structural clusters don't contain protease inhibitors, so we used ProteinCartography semantic analysis of protein annotations to identify high- and low-confidence clusters of protease inhibitors. High-confidence clusters are structural clusters with within-cluster TM-scores > 0.2 **[18]** and annotations clearly related to protease inhibitor function. Low-confidence clusters fail to satisfy one of these requirements, but have some protease inhibitor-related annotations and therefore contain potential genes of interest. We retained all low-confidence cluster members with the caveat that many of them may not be true protease inhibitors.

We then used our NovelTree **[19]** analysis of chelicerate gene families **[12]** to identify orthogroups of secreted proteins putatively involved in protease inhibition within these structural clusters. For orthogroups with high-quality protease inhibitor annotations, we retained orthogroups with more than five members with DeepSig secretion signals, and for which secreted proteins comprised more than 5% of the orthogroup. For orthogroups with low-quality protease inhibitor annotations, we were more stringent and required that ≥ 25% of the orthogroup had DeepSig secretion signals, and that this represented more than 25 total proteins. We also required that ≥ 50% of orthogroup members were annotated as protease inhibitors or secreted PUFs. We moved forward with the 36 orthogroups that fit these requirements (14 high-quality, 22 low-quality).

> Check out "Notebook 02" for **code to prepare metadata files**, "Notebook 03" for **ProteinCartography analysis code**, and the ProteinCartography config file in our GitHub repo.

Phylogenetic profiling of tick orthogroups

> **Note**
>
> The methods described below represent a legacy approach that we used as we were completing the development of our phylogenetic profiling pipeline. The proteome-wide analysis and approach found in **[12]** represents a distinct, and more mature implementation of these concepts. This smaller-scale analysis here served only to let us quickly rank gene families in this case study.

In parallel work **[12]**, we calculated gene copy number and curated species-level, parasitism-related trait data across 40 chelicerates that vary in their ability to suppress host detection. In this work, we used that data to help identify which of our 36 orthogroups of interest most strongly predict the host detection suppression trait. First, we normalized the distributions of species gene copy number by $\log_{10}$-transforming them [$\log_{10}(x+1)$, adding one to avoid undefined values for $\log_{10}(0)$]. To account for the statistical non-independence of species and their phenotypes as induced by their shared evolutionary history **[20]**, we applied a phylogenetic generalized least-squares transformation **[21]** to both the species trait and gene copy number data using the species tree inferred from SpeciesRax **[22]** (applied to data within the `fit_lasso_counts` function defined in "trait_mapping_functions.R"). This transformation effectively "regresses out" the effect of evolutionary non-independence on trait variation, returning the "residual" trait variation not explained by common ancestry alone.

Using these transformed data, we subsequently conducted Lasso regression (i.e., L1 regularization) to predict whether species suppress host detection using gene copy number, implemented using the R package "familiar" **[23]** in a custom function (`fit_lasso_counts`). Specifically, we predicted detection suppression as a binomial outcome, using 10-fold cross-validation with feature selection using the `lasso_binomial` method, clustering features prior to feature selection using the `hclust` method. Using this analysis, we selected OG0000058, as it had the highest model variable importance for which the partial dependence plots also demonstrated a positive relationship between copy number and probability of supressing host detection.

Check out our phylogenetic profiling code.

Selecting tick candidates for protein–protein interaction (PPI) prediction

Our next step was to take members of our top-scoring protease inhibitor orthogroup OG0000058 and use protein–protein interaction prediction tools to predict their targets in the human proteome. For this case study, we chose to follow up on 10 members OG0000058 encoded by the lone star tick *Amblyomma americanum*. *A. americanum* has 32 genes in this gene family, so we filtered down to choose the ones most likely to interact with the host. We picked 10 genes with secretion signals and the highest expression level in the female tick salivary gland, according to our differential expression analysis of *A. americanum* **[24]**. For analyses where we only used a single protease inhibitor, we used Amblyomma-americanum_evm.model.contig-94090-1.4 as a representative, since it had one of the highest-quality structures of the group (pLDDT = 83.8).

For more **details on how we selected candidate protease inhibitors**, see "Notebook 05" in our GitHub repo.

Selecting human targets for protein–protein interaction (PPI) prediction

OG0000058 is a family of trypsin-inhibitor-like (TIL) domain serine protease inhibitors, so we decided to test for interactions against serine proteases. To select our targets, we accessed protein sequences and structures for proteins with annotations related to "serine protease" from the UniProt human reference proteome (accession UP000005640). We then filtered out all proteins > 700 amino acids in length. We chose this cutoff because D-SCRIPT can't reliably predict interactions for protein complexes longer than a combined length of 1,000 amino acids (AAs), and in our initial

tests of AF-Multimer, performance dropped off for protein complexes longer than ~700 AAs.

Running AF-Multimer

AF-Multimer **[11]** takes in combined FASTA files of bait and target proteins to fold into a complex. To prepare the combined FASTA files separated by a semicolon for AF-Multimer predictions, we created two scripts using Biopython (v1.81) **[25]**: "make-afmultimer-fasta-from-seqs.py" to make combinations from two FASTA files and "make-afmultimer-from-tsv.py" to make combinations from a TSV list. We used ColabFold (v1.5.2) **[26]** to launch the AlphaFold2-Batch notebook **[27]** on Google Colab to make AF-Multimer complex predictions **[11]**. For all runs, we set the `msa_mode` parameter to use MMseqs2 **[28]** using the UniRef + Environmental database **[26]**, the `num_models` parameter to five models per prediction, and the `num_recycles` parameter to three.

> Access **AF-Multimer inputs and raw outputs** on Zenodo.

AF-Multimer analysis and metrics

To analyze the AF-Multimer predictions, we used the LazyAF analysis notebook **[29]** to take the highest-ranked prediction and calculate the predicted template modeling (pTM) score, the interface predicted template modeling (ipTM) score, and the ranking confidence score. The ranking confidence score equals $0.2 \times \mathrm{pTM} + 0.8 \times \mathrm{ipTM}$, and a high-confidence score is at or above 0.75. We also modified the existing LazyAF notebook to add a calculation for the average pLDDT of the predicted complex structure, which is available here. To calculate additional protein complex metrics, such as the predicted DockQ (pDockQ) of the interacting residues at the binding interface, we used the "AF2multimer-analysis" script from the predictomes.org toolkit **[30]**. We considered a high-confidence interaction score based on pDockQ to be at or above a threshold of 0.5 **[31][32]**.

See “Notebook 06” in our GitHub repo for **processing AF-Multimer raw results**.

Additional metrics, such as each predicted complex’s pLDDT and predicted alignment error (PAE), are in the “2024-04-24-final-serine-protease-filtered-interface-comps-results.tsv” file on GitHub.

Running D-SCRIPT

We ran D-SCRIPT (v0.2.6) using the topsy-turvy model of human–human PPIs, as recommended in the documentation **[10]**. We automated the steps for configuring inputs and running predictions for D-SCRIPT with the script “make_dscript_PPI_predictions.py.” A high-confidence PPI interaction score with D-SCRIPT is at or above 0.5.

See “Notebook 07” in our GitHub repo for **comparison of D-SCRIPT and AF-Multimer results**.

Protease cell and tissue type expression analysis

Finally, we wanted to determine which tissues and cell types in which the predicted targets of tick protease inhibitors are expressed. We obtained the predicted expression of each human serine protease hit from the NCBI Gene Database using an RNA-seq study of 95 humans from 27 tissues **[33]**. We determined a protease to be expressed in the skin if it was labeled as being expressed in “skin” and assigned proteases to be putatively expressed in immune cells if they were highly expressed in lymphoid tissues such as bone marrow and spleen. We also labeled chymase (P23946) and tryptase (Q15661) as being produced by immune cells since they originate from mast cells (tissue-resident granulated immune cells). We then followed up on our predicted skin and immune cell-expressed proteases using the Human

Protein Atlas to identify specific cell types responsible for producing the protease. We performed data analysis, parsing, and plotting in R (v4.3.3) using tidyverse (v2.0.0) **[34]** and patchwork (v1.2.0) **[35]** packages.

> View expression data for the 35 proteases of interest and see "Notebook 08" for our **analysis**.

Additional methods

We used ChatGPT to help write and clean up code. We also provided ChatGPT with starting text to reorganize into one of our pub templates. ChatGPT and Grammarly Premium also suggested wording ideas, and then we chose which small phrases or sentence structure ideas to use.

We used arcadia-pycolor (v0.5.1) **[36]** and arcadiathemeR (v0.1.1) **[37]** to generate figures before manual adjustment.

> All **code** we generated and used for the pub is available in this GitHub repository (DOI: 10.5281/zenodo.15243681), which includes code for ProteinCartography results analysis, phylogenetic profiling, generating PPI prediction FASTA file inputs, D-SCRIPT runs, and AF-Multimer analyses.

The results

SHOW ME THE DATA

> Protein sequences, annotations, and structures are on Zenodo. We've also uploaded our full ProteinCartography analysis and our raw AF-Multimer results.
>
> Phylogenetic profiling results, D-SCRIPT results, and AF-Multimer summary results can be found in our GitHub repo.

We're interested in scalably predicting the targets of tick salivary effectors. As a case study, we decided to test new methodology to do this for tick protease inhibitors. Protease inhibitors are abundant in tick saliva, and other groups have successfully used PPI prediction tools to identify the host protease targets of parasite protease inhibitors **[38]**.

Structural clustering identifies 36 gene families of putative protease inhibitors

The first step in this case study was to identify protease inhibitor families that ticks may use to interact with their hosts. To do this, we used sequence-based annotations of 15 tick proteomes to select all proteins predicted to be involved in protease inhibition. This comprised 3,453 putative protease inhibitors. Because many tick proteins don't receive high-confidence functional annotations and structural comparisons are more robust to sequence divergence than sequence-based annotation methods, we decided to use structure-based methods to identify unannotated protease inhibitors. To this end, we used our in-house tool ProteinCartography **[17]** to define major structural clusters of known protease inhibitors and secreted proteins of unknown function (PUFs). We clustered 3,453 annotated tick protease inhibitors and 12,041 secreted PUFs with a database of 7,008 venom toxins, which includes diverse protease inhibitors from different venomous species. This produced 32 Leiden clusters. By looking through the ProteinCartography cluster TM-scores (a measure of intra-cluster structural similarity) and ProteinCartography cluster semantic analysis (aggregated functional annotations), we identified seven high-confidence protease inhibitor Leiden clusters (LC06, LC13, LC14, LC15, LC20, LC22, LC29) and four low-confidence protease inhibitor Leiden clusters (LC02, LC05, LC17, LC21) (Figure 1).

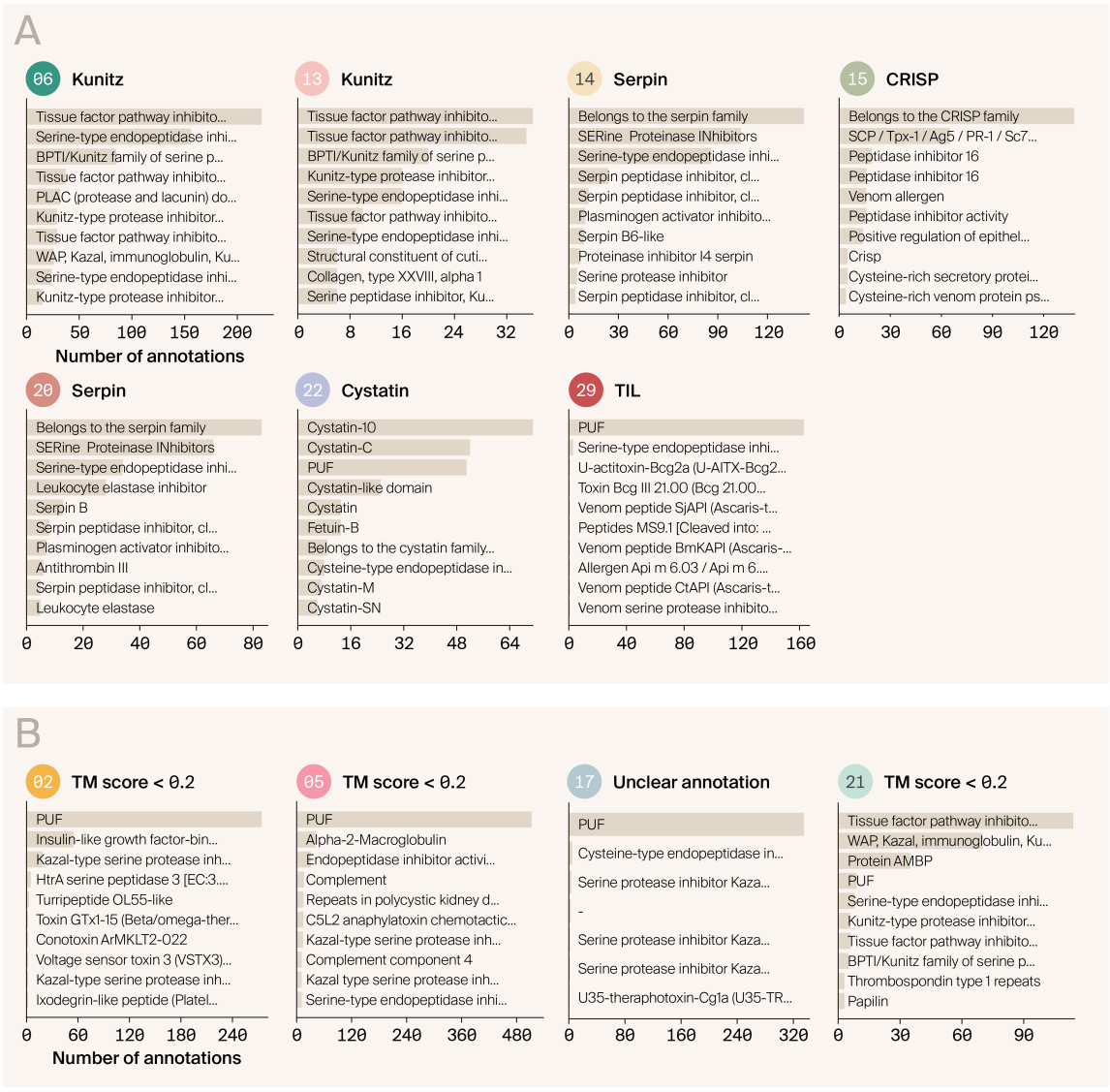

Figure 1

Semantic analysis of A) seven high-confidence protease inhibitor clusters and B) four low-confidence structural clusters.

High confidence clusters (LC06, LC13, LC14, LC15, LC20, LC22, LC29) have a within-cluster TM-score > 0.2, and annotations related to protease inhibition. We have labeled the seven high-confidence clusters with their consensus domain annotations at the top of their plot. Low confidence clusters (LC02, LC05, LC21) with low within-cluster TM-score are labeled at the top of their plot. We consider LC17 to be a low-confidence cluster despite having a within-cluster TM-score > 0.2 because it has inconclusive annotations.

For our downstream analyses, we needed to work with individual gene families, not large clusters of proteins. As we'd defined tick gene families using NovelTree **[19]** in a parallel effort **[12]**, we were able to identify the tick gene families that were present in our protease inhibitor structural clusters. Overall, this gave us a total of 36 gene families of putative protease inhibitors, 14 coming from high-confidence structural clusters and 22 coming from low-confidence clusters. We chose to retain gene families originating from low-confidence clusters in an effort to be comprehensive, with the caveat that this choice could introduce gene families that didn't have protease inhibitor function into our downstream analyses.

> You can find the **full ProteinCartography results** in "tick_PUFs_PIs_1200_plus_toxinDB_carto_run3.zip" on Zenodo.

OG0000058 is a family of TIL-domain protease inhibitors predicted to suppress host detection

Next, we had to prioritize a single family of protease inhibitors to test in our target-prediction framework. While we were doing this target-prediction pilot, we were also kicking off a larger effort to use phylogenetic profiling of ticks and their relatives (collectively called chelicerates) to identify the genetic basis by which ticks and other parasites block host detection mechanisms like itch, pain, and inflammation **[12]**. From this larger effort, we'd inferred the evolutionary histories (patterns of duplication, transfer, and loss) for all tick gene families, including our 36 gene families of interest. We used our evolutionary data to rank our protease inhibitor families based on how likely they are to be associated with suppressing host detection. We reasoned that protease inhibitors linked to suppressing host detection would be more likely to target host proteases than endogenous tick proteases.

We found that the top three gene families with the highest model variable importance (most likely to suppress host detection based on our phylogenetic analysis) were OG0001324, OG0000480, and OG0000058 (Figure 2, A). When we looked at the annotation for each group, we found that OG0001324 derived from a low-confidence structural cluster and doesn't actually contain protease inhibitors. Further analysis showed that they're a family of secreted proteins with some similarity to IL-17-like cytokines. While these putative IL-17-like proteins are very interesting to us (and we'll

release a pub on them soon!), we still wanted to try our target prediction pipeline with protease inhibitors, so we turned to the second- and third-ranked hits. OG0000480 is annotated as alpha-2 macroglobulin-like protease inhibitors, and OG0000058 is annotated as a family of trypsin-inhibitor-like (TIL) domain protease inhibitors. OG0000480 is negatively associated with the detection-suppression trait, whereas OG0000058 is positively associated with host detection suppression (Figure 2, B). So, we chose to use OG0000058 as our candidate family of host-directed protease inhibitors. TIL domain protease inhibitors are predicted to inhibit serine proteases **[39]**, but their role in tick biology is relatively unknown.

Our **phylogenetic profiling results** are available on GitHub.

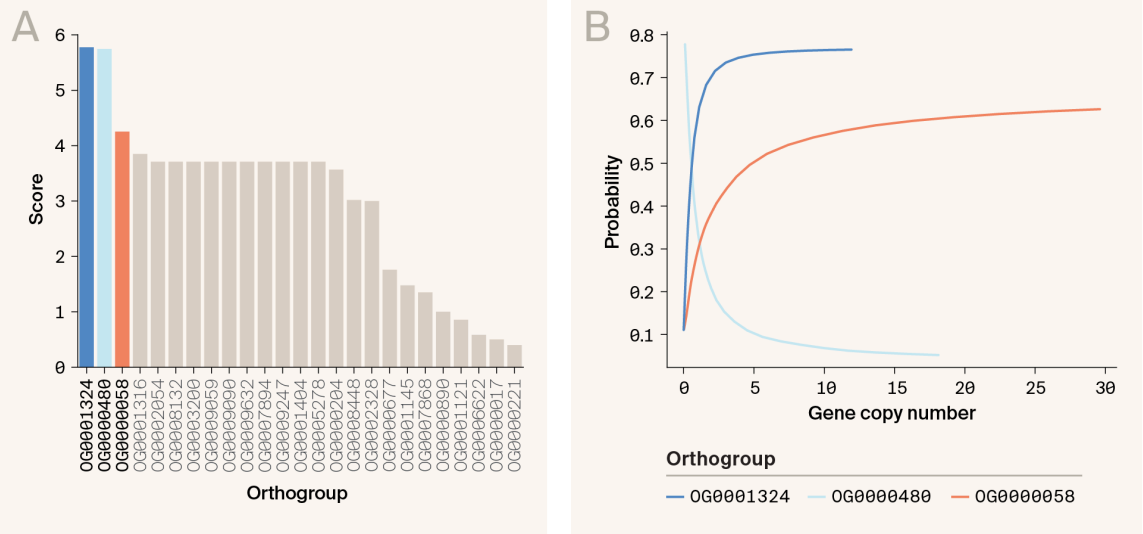

Figure 2

Phylogenetic profiling of putative tick protease inhibitor orthogroups.

(A) Model variable importance (Borda scores aggregated across cross-validation replicates) of different tick orthogroups in predicting the detection-suppression trait.

(B) Partial dependence plots for the top orthogroups OG0001324, OG0000480, and OG0000058 in the validation phase of model construction, depicting how, in the fitted model, the probability of host detection suppression increases or decreases along with increasing copy number.

AF-Multimer predicts 34 targets for a single representative of OG0000058

We decided to test the ability of two different tools to predict tick-human protein–protein interactions (PPIs): D-SCRIPT **[10]** and AF-Multimer **[11]**. D-SCRIPT is a deep learning tool for predicting PPIs based on sequence inputs, and AF-Multimer is a structure-based approach that predicts PPIs by folding multiple proteins together in a complex, which can then be scored and ranked. D-SCRIPT is much cheaper and faster to run, but AF-Multimer has been demonstrated to perform well in predicting cross-species PPIs, specifically for protease inhibitors **[38]**. We began by screening a single representative of OG0000058 from the lone star tick *Amblyomma americanum* (Amblyomma-americanum_evm.model.contig-94090-1.4) against 527 human serine proteases using D-SCRIPT and AF-Multimer to directly compare and evaluate their performance (Figure 3). While we got high-confidence hits from both tools, there was little overlap between them. We suspect the discordance between D-SCRIPT and AF-Multimer is because the PPI model underlying D-SCRIPT is built based on human–human PPI predictions and doesn't generalize well to cross-species PPI interactions. Given this, we decided to move forward with just the hits from AF-Multimer. This analysis identified 34 reviewed UniProt accessions with an AF-Multimer ranking confidence score above 0.75 (see the full results, including unreviewed accessions here).

Question

What other tools should we try?

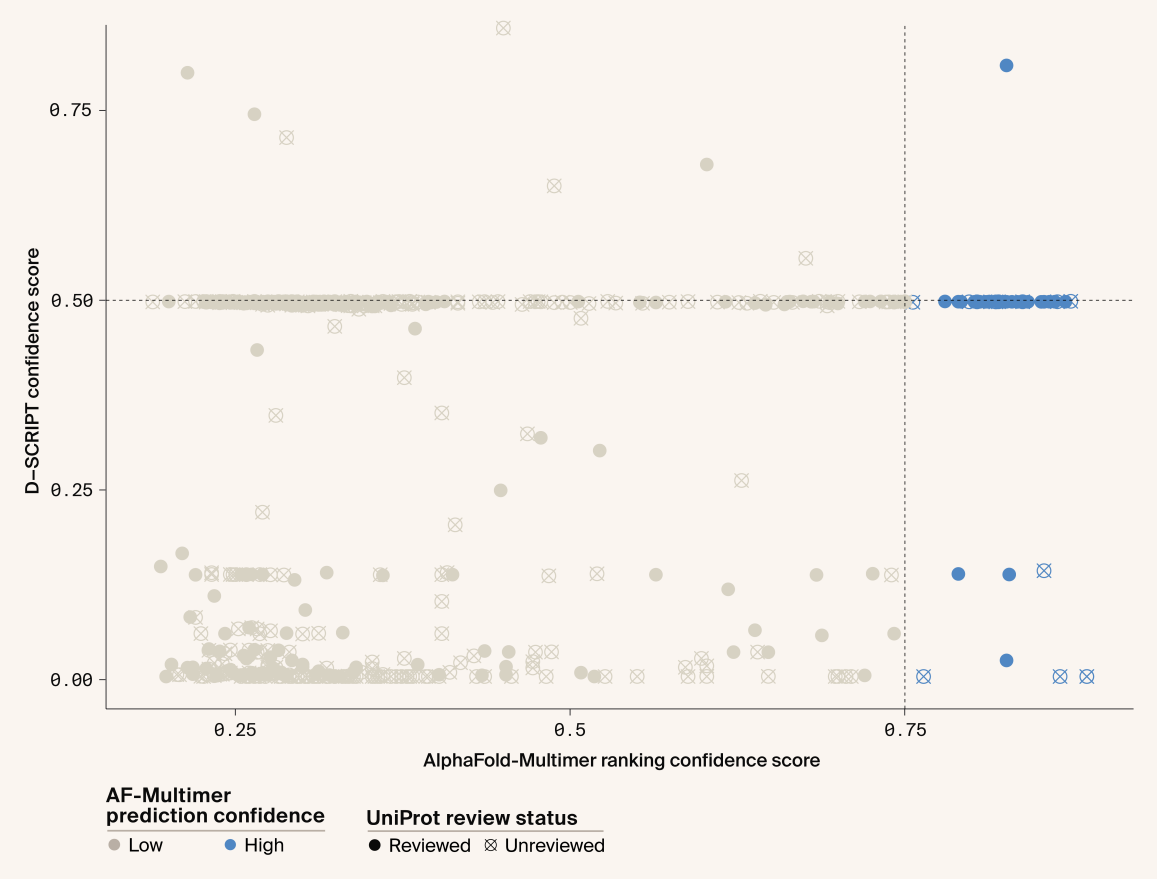

Figure 3

Comparisons of PPI predictions between sequence- (D-SCRIPT) and structure-based (AF-Multimer) tools.

Plot shows tick–human PPI predictions for protease inhibitor 94090-1.4 against 527 human serine proteases. Confident hits from D-SCRIPT are at or above 0.5 and confident hits from AF-Multimer are at or above 0.75, indicated with dashed lines. Filled-in points are reviewed UniProt accessions, whereas crossed-through points are unreviewed UniProt accessions, which are primarily isoforms of the reviewed accessions. Blue points indicate that the prediction is a high-confidence AF-Multimer hit — these are the potential targets against which we decided to test our nine remaining tick protease inhibitors.

Predicted targeted proteases are commonly expressed in the skin, immune cells, or the pancreas

To understand the physiological relevance of these potential targets, we looked at tissue expression data for each human protease (see all our hits with tissue expression data here). Excitingly, 14 of these proteases are primarily expressed in immune cells or the skin, and would be plausible targets of tick protease inhibitors. Ticks interact intimately with human skin and would need to counteract any host proteases produced by skin cells or local immune cells that are involved in inflammation, itch, pain, or other modes of host defense. Furthermore, of these skin- or immune-cell-expressed proteases, several are implicated in various dermatological pathologies characterized by itch, pain, and inflammation (Table 1), highlighting the importance of regulating the activity levels of these enzymes. However, it's clear that not all of the hits are relevant to tick biology. For example, 12 highly scoring hits were to pancreas-produced proteases, which the tick is unlikely to encounter. It's possible that these 34 hits are all genuine potential interactors for the tick protease inhibitor (irrespective of whether the tick would actually encounter them in skin). It's also possible that this large number of physiologically implausible hits reflects a high false-positive rate.

> **Question**
>
> What do you think? Does a single protease inhibitor really interact with all these different proteases, or does this reflect a high false-positive rate?

Annotation	UniProt name	UniProt ID	Source	Associated skin diseases
Azurocidin	CAP7	P20160	Immune cells (neutrophils)	
Cathepsin G	CATG	P08311	Immune cells (neutrophils, B cells)	Atopic dermatitis **[40]**, generalized pustular psoriasis **[41] [42]**, psoriasis **[43][44][45]**
Chymase	CMA1	P23946	Immune cells (mast cells)	Atopic dermatitis **[46][47][48][49] [50]**
Granzyme B	GRAB	P10144	Immune cells (NK cells, cytotoxic T cells)	Atopic dermatitis **[51]**, psoriasis, various autoimmune skin diseases **[52][53] [54][55]**
Granzyme H	GRAH	P20718	Immune cells (NK cells, cytotoxic T cells)	
Granzyme M	GRAM	P51124	Immune cells (NK cells, cytotoxic T cells)	
Myeloblastin	PRTN3	P24158	Immune cells (polymorphonuclear leukocytes, neutrophils)	
Neutrophil elastase	ELNE	P08246	Immune cells (neutrophils)	Bullous pemphigoid **[56]**, atopic dermatitis **[57]**, psoriasis **[44] [45]**
Serine protease 33	PRS33	Q8NF86	Immune cells (macrophages)	
Tryptase alpha/beta-1	TRYB1	Q15661	Immune cells (mast cells)	Atopic dermatitis **[58]**, mastocytosis **[59]**

Kallikrein-7	KLK7	P49862	Skin (keratinocytes)	Netherton's syndrome, atopic dermatitis **[60][61][62][63]**
Kallikrein-8	KLK8	O60259	Skin (keratinocytes)	Various skin diseases **[64] [65]**
Kallikrein-9	KLK9	Q9UKQ9	Skin (keratinocytes)	
Kallikrein-14	KLK14	Q9P0G3	Skin	Netherton's syndrome **[66]**, various skin diseases **[65]**

Table 1

Immune- and skin-expressed serine proteases predicted to be targeted by the tick protease inhibitor Amblyomma-americanum_evm.model.contig-94090-1.4.

These human serine proteases are predominantly expressed in skin or immune cells according to the Human Protein Atlas, and we predict that they're inhibited by protease inhibitor Amblyomma-americanum_evm.model.contig-94090-1.4. Where possible, we also indicate the cell type that the Human Protein Atlas predicts will produce the protease. Many of these proteases are associated with skin diseases.

Incorporating pDockQ score didn't substantially change our results

In an attempt to separate any potential false positives from the true positives, we rescored our hits using predicted DockQ (pDockQ) score. While the AF-Multimer ranking confidence score is a combined metric of the interface score (ipTM) and an overall structural score of the complex (pTM), pDockQ score is the average quality of interacting residues of the complex **[31][32]**. We hoped this additional metric would give us more information on which hits to consider high-confidence. However, only four of the 34 hits fell below the pDockQ confidence threshold of 0.5. Notably, three out of the four hits that were excluded by the pDockQ score were produced by immune cells

and were hits we'd initially considered to be plausible targets of tick protease inhibitors (myeloblastin: pDockQ score of 0.464, granzyme B: pDockQ score of 0.467, tryptase alpha/beta-1: pDockQ score of 0.469). Our file with full results, including pDockQ and AF-Multimer scores, is available here.

> **Question**
>
> Are there other metrics or cutoffs that we should look at that could help give confidence in our hits, and eliminate false positives?

Next, we compared the scores of each protease with pDockQ and AF-Multimer confidence. Initially, we thought higher-probability hits might consistently receive high scores from both metrics, and low-probability hits would be noisier. However, both scoring systems gave different top-ranked hits, and there wasn't an obvious way for us to use the scores in conjunction (Figure 4). pDockQ gave more biologically plausible top hits (neutrophil elastase, complement factor D, and mast cell chymase, all related to the innate immune response) compared to the AF-Multimer confidence score (pancreatic chymotrypsins, related to digestion). However, the fourth-ranked hit for pDockQ is another pancreatic chymotrypsin.

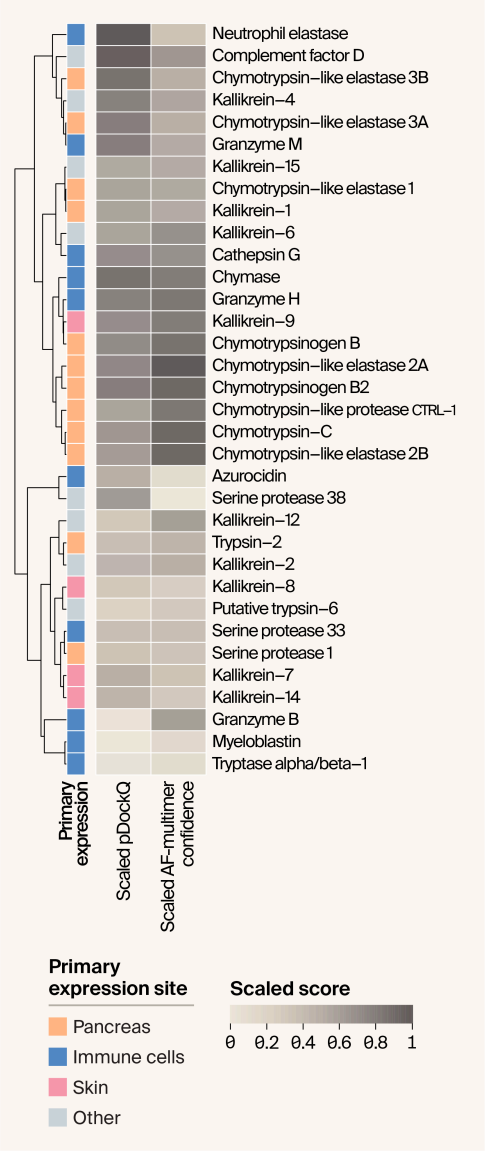

Figure 4

Comparison of pDockQ and AF-Multimer confidence rankings for putative targets of Amblyomma-americanum_evm.model.contig-94090-1.4.

Heatmap of pDockQ scores (left) and AF-Multimer confidence (right) for each of the 34 predicted targets of Amblyomma-

americanum_evm.model.contig-94090-1.4, clustered by Euclidean distance. We have scaled these scores to make the relative rankings of each hit comparable between the two scoring systems. Darker shades correspond to higher ranking. We've color-coded proteases produced by skin cells, immune cells, or the pancreas.

Other members of OG0000058 are predicted to have similar target profiles

As another approach to enrich for biologically plausible hits, we decided to predict targets for other members of the OG0000058 family. This gene family is highly expanded in ticks, with 32 distinct copies in *A. americanum* alone. We selected nine other secreted, salivary-gland-expressed members of this family from *A. americanum* to predict protein–protein interactions. Ideally, we'd perform the same initial search against all human serine proteases for nine other representatives, and compare if they have the same or different targets. However, this wasn't feasible because of time and computational resource limitations, so we instead used the 34 positive hits from the initial screen to make predictions for the remaining nine protease inhibitors. We weren't sure if we should expect them all to have similar target profiles, or if their expansion in ticks reflects a diversification of targets. Regardless, we reasoned that any false positives from our initial search with one member might consistently score poorly when tested against more family members.

In doing this, we found that in most cases, the other members of the OG0000058 family also had high-confidence interactions predicted for the 34 proteases from the initial search (Figure 5). We scored these interactions with pDockQ (Figure 5, A) and AF-Multimer confidence (Figure 5, B), but both scoring systems overall produced similar results. Using either metric, the majority of the interactions were predicted to be high confidence (> 0.75 for AF-Multimer confidence and > 0.5 for pDockQ). The clearest signal we saw from this analysis was that the protease inhibitor Amblyomma-americanum_evm.model.contig-8661-1.1 had below-threshold scores for every

predicted interaction. When we looked into it further, we found that this protein had a very low-quality structure (pLDDT = 36.9 out of 100), so we attribute these disparate results to this low structural quality. Our file with results for all 10 protease inhibitors is available here.

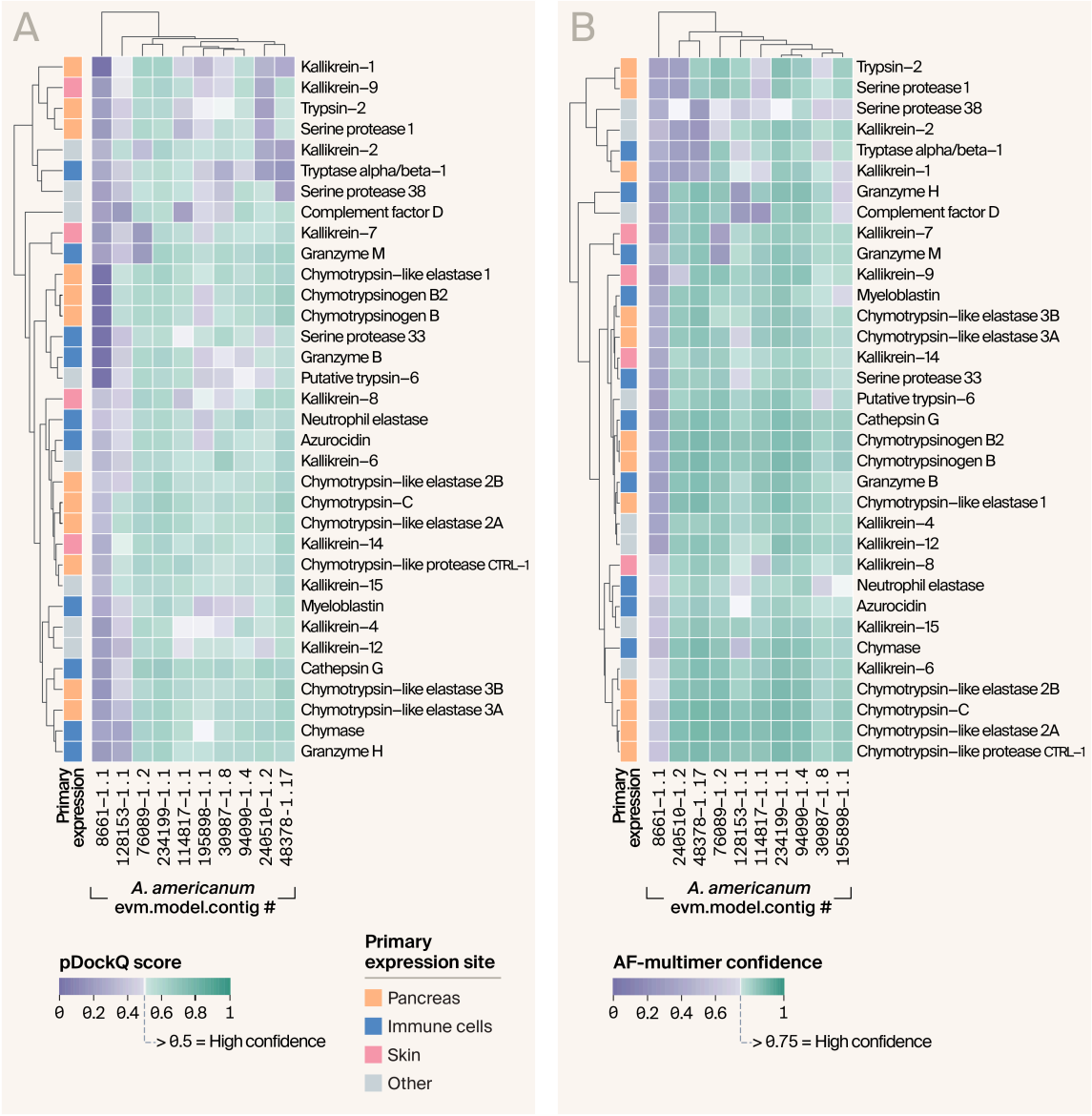

Figure 5

Predicted protease targets of ten members of OG0000058 from *Amblyomma americanum*.

Heatmap of pDockQ scores (A) and AF-Multimer confidence (B) for the 34 initial predicted targets of Amblyomma-americanum_evm.model.contig-94090-1.4, tested against nine other family members. High-confidence hits are green (A: pDockQ > 0.5, B: AF-Multimer confidence > 0.75), whereas low-confidence scores are purple. We've clustered the targets and the protease inhibitors by Euclidean distance and have color-coded proteases produced by skin cells, immune cells, or the pancreas.

Question

Do these family members really all target similar proteases? Or is this a sign that our approach isn't giving us sufficient resolution?

Conclusion

We were able to use AF-Multimer to predict targets of a family of TIL-domain tick protease inhibitors, but we have no idea how much confidence we should have in these hits. Some of these targets seem reasonable (i.e., produced by skin or immune cells), but we found an almost equal number of physiologically irrelevant targets (i.e., pancreatic enzymes). These *could* represent real potential interactors, as there's precedent for tick protease inhibitors to have the capacity to inhibit proteases not found in the skin. For example, serpin Rms-3 from the tick *Rhipicephalus microplus* can inhibit pancreatic enzymes chymotrypsin and elastase **[67]**. However, it's also possible that many of our computational hits are false positives, and in reality, these TIL-domain protease inhibitors are active against a much smaller range of targets. Without knowing how to distinguish between these two possibilities, it's hard to justify moving the project forward.

What's next?

We think there's a lot of promise in predicting the targets of tick effector proteins, and are still excited about what this approach can teach us about useful ways to manipulate skin biology. We aren't sold on this particular method though, given that we don't have a great way to evaluate our true vs. false positive rates.

Right now, we think the right thing to do would be to benchmark our computational search against actual experimental data. However, we don't know of any datasets that would be right to use as the "ground truth" for benchmarking our pipeline. While some targets of some tick protease inhibitors have been identified, we don't know of any studies that have comprehensively evaluated the activity of a single tick protease against all human proteases (the experimental equivalent of our computational search).

Overall, we're still interested in exploring alternative or additional approaches to solve this problem, but we've gotten stuck. We're sharing this case study in case others in the community have opinions or feedback on what we should try next.

References

1. Nuttall PA. (2019). Wonders of tick saliva. https://doi.org/10.1016/j.ttbdis.2018.11.005
2. Ticks. (2023). https://doi.org/10.1016/c2021-0-00008-0
3. Ujiie H, Rosmarin D, Schön MP, Ständer S, Boch K, Metz M, Maurer M, Thaci D, Schmidt E, Cole C, Amber KT, Didona D, Hertl M, Recke A, Graßhoff H, Hackel A, Schumann A, Riemekasten G, Bieber K, Sprow G, Dan J, Zillikens D, Sezin T, Christiano AM, Wolk K, Sabat R, Kridin K, Werth VP, Ludwig RJ. (2022). Unmet Medical Needs in Chronic, Non-communicable Inflammatory Skin Diseases. https://doi.org/10.3389/fmed.2022.875492
4. Ribeiro JMC, Mans BJ. (2020). TickSialoFam (TSFam): A Database That Helps to Classify Tick Salivary Proteins, a Review on Tick Salivary Protein Function and Evolution, With Considerations on the Tick Sialome Switching Phenomenon. https://doi.org/10.3389/fcimb.2020.00374
5. Walsh PN, Ahmad SS. (2002). Proteases in blood clotting. https://doi.org/10.1042/bse0380095
6. Moali C, Hulmes DJ. (2009). Extracellular and cell surface proteases in wound healing: new players are still emerging. https://doi.org/10.1684/ejd.2009.0770
7. Heutinck KM, ten Berge IJM, Hack CE, Hamann J, Rowshani AT. (2010). Serine proteases of the human immune system in health and disease. https://doi.org/10.1016/j.molimm.2010.04.020
8. Akiyama T, Lerner EA, Carstens E. (2015). Protease-Activated Receptors and Itch. https://doi.org/10.1007/978-3-662-44605-8_13
9. McDougall JJ, Muley MM. (2015). The Role of Proteases in Pain. https://doi.org/10.1007/978-3-662-46450-2_12

10 Sledzieski S, Singh R, Cowen L, Berger B. (2021). D-SCRIPT translates genome to phenome with sequence-based, structure-aware, genome-scale predictions of protein-protein interactions. https://doi.org/10.1016/j.cels.2021.08.010

11 Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, Žídek A, Bates R, Blackwell S, Yim J, Ronneberger O, Bodenstein S, Zielinski M, Bridgland A, Potapenko A, Cowie A, Tunyasuvunakool K, Jain R, Clancy E, Kohli P, Jumper J, Hassabis D. (2021). Protein complex prediction with AlphaFold-Multimer. https://doi.org/10.1101/2021.10.04.463034

12 Borges AL, Chou S, Patton AH, Reiter T, Weiss ECP, York R. (2025). Comparative phylogenomic analysis of chelicerates points to gene families associated with long-term suppression of host detection. https://doi.org/10.57844/ARCADIA-4E3B-BBEA

13 Savojardo C, Martelli PL, Fariselli P, Casadio R. (2017). DeepSig: deep learning improves signal peptide detection in proteins. https://doi.org/10.1093/bioinformatics/btx818

14 Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, Ogata H. (2019). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. https://doi.org/10.1093/bioinformatics/btz859

15 Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. https://doi.org/10.1093/molbev/msab293

16 Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, dos Santos Costa A, Fazel-Zarandi M, Sercu T, Candido S, Rives A. (2023). Evolutionary-scale prediction of atomic-level protein structure with a language model. https://doi.org/10.1126/science.ade2574

17 Avasthi P, Bigge BM, Celebi FM, Cheveralls K, Gehring J, McGeever E, Mishne G, Radkov A, Sun DA. (2024). ProteinCartography: Comparing proteins with structure-based maps for interactive exploration. https://doi.org/10.57844/ARCADIA-A5A6-1068

18 Xu J, Zhang Y. (2010). How significant is a protein structure similarity with TM-score = 0.5? https://doi.org/10.1093/bioinformatics/btq066

19 Celebi FM, Chou S, McGeever E, Patton AH, York R. (2024). NovelTree: Highly parallelized phylogenomic inference. https://doi.org/10.57844/ARCADIA-Z08X-V798

20 Felsenstein J. (1985). Phylogenies and the Comparative Method. https://doi.org/10.1086/284325

21 Butler MA, Schoener TW, Losos JB. (2000). THE RELATIONSHIP BETWEEN SEXUAL SIZE DIMORPHISM AND HABITAT USE IN GREATER ANTILLEAN*ANOLIS*LIZARDS. https://doi.org/10.1111/j.0014-3820.2000.tb00026.x

22 Morel B, Schade P, Lutteropp S, Williams TA, Szöllősi GJ, Stamatakis A. (2022). SpeciesRax: A Tool for Maximum Likelihood Species Tree Inference from Gene Family Trees under Duplication, Transfer, and Loss. https://doi.org/10.1093/molbev/msab365

23 Zwanenburg A, Löck S. (2021). familiar. https://github.com/alexzwanenburg/familiar

24 Borges AL, Reiter T, Weiss ECP. (2025). An interactive visualization tool for Amblyomma americanum differential expression data. https://doi.org/10.57844/ARCADIA-6833-4117

25 Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJL. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. https://doi.org/10.1093/bioinformatics/btp163

26 Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. (2022). ColabFold: making protein folding accessible to all. https://doi.org/10.1038/s41592-022-01488-1

27 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. (2021). Highly accurate protein structure prediction with AlphaFold. https://doi.org/10.1038/s41586-021-03819-2

28 Steinegger M, Söding J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. https://doi.org/10.1038/nbt.3988

29 McLean TC. (2024). LazyAF, a pipeline for accessible medium-scale*in silico*prediction of protein-protein interactions. https://doi.org/10.1101/2024.01.29.577767

30 Schmid EW, Walter JC. (2024). Predictomes: A classifier-curated database of AlphaFold-modeled protein-protein interactions. https://doi.org/10.1101/2024.04.09.588596

31 Zhu W, Shenoy A, Kundrotas P, Elofsson A. (2023). Evaluation of AlphaFold-Multimer prediction on multi-chain protein complexes. https://doi.org/10.1093/bioinformatics/btad424

32 Hou Y, Xie T, He L, Tao L, Huang J. (2023). Topological links in predicted protein complex structures reveal limitations of AlphaFold. https://doi.org/10.1038/s42003-023-05489-4

33 Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjöstedt E, Lundberg E, Szigyarto CA-K, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson Å, von Feilitzen K, Forsberg M, Zwahlen M, Olsson I, Navani S, Huss M, Nielsen J, Ponten F, Uhlén M. (2014). Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. https://doi.org/10.1074/mcp.m113.035600

34 Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. (2019). Welcome to the Tidyverse. https://doi.org/10.21105/joss.01686

35 Pedersen, Thomas Lin. (2024). patchwork. %7Bhttps://github.com/thomasp85/patchwork%7D

36 arcadia-pycolor. (2025). https://github.com/Arcadia-Science/arcadia-pycolor

37 arcadiathemeR. (2024). https://github.com/Arcadia-Science/arcadiathemeR

38 Homma F, Huang J, van der Hoorn RAL. (2023). AlphaFold-Multimer predicts cross-kingdom interactions at the plant-pathogen interface. https://doi.org/10.1038/s41467-023-41721-9

39 Chmelař J, Kotál J, Langhansová H, Kotsyfakis M. (2017). Protease Inhibitors in Tick Saliva: The Role of Serpins and Cystatins in Tick-host-Pathogen Interaction. https://doi.org/10.3389/fcimb.2017.00216

40 Breuckmann F, von Kobyletzki G, Avermaete A, Kreuter A, Altmeyer P, Gambichler T. (2002). Modulation of cathepsin G expression in severe atopic dermatitis following medium-dose UVA1 phototherapy. https://doi.org/10.1186/1471-5945-2-12

41 Liu Y, Li H, Meng S, Xu Y, Ni S, Sun Y, Zhou W. (2023). Newly revealed variants of SERPINA3 in generalized pustular psoriasis attenuate inhibition of ACT on cathepsin G. https://doi.org/10.1038/s10038-023-01139-z

42 Frey S, Sticht H, Wilsmann-Theis D, Gerschütz A, Wolf K, Löhr S, Haskamp S, Frey B, Hahn M, Ekici AB, Uebe S, Thiel C, Reis A, Burkhardt H, Behrens F, Köhm M, Rech J, Schett G, Assmann G, Kingo K, Kõks S, Mössner R, Prinz JC, Oji V, Schulz P, Muñoz LE, Kremer AE, Wenzel J, Hüffmeier U. (2020). Rare Loss-of-Function Mutation in SERPINA3 in Generalized Pustular Psoriasis. https://doi.org/10.1016/j.jid.2019.11.024

43 Guo J, Tu J, Hu Y, Song G, Yin Z. (2019). Cathepsin G cleaves and activates IL-36γ and promotes the inflammation of psoriasis https://doi.org/10.2147/dddt.s194765

44 Krasavin MYu, Gureev MA, Garabadzhiu AV, Pashkin AYu, Zhukov AS, Khairutdinov VR, Samtsov AV, Shvets VI. (2019). Inhibition of Neutrophil Elastase and Cathepsin G As a New Approach to the Treatment of Psoriasis: From Fundamental Biology to Development of New Target-Specific Drugs. https://doi.org/10.1134/s1607672919040082

45 Sullivan GP, Henry CM, Clancy DM, Mametnabiev T, Belotcerkovskaya E, Davidovich P, Sura-Trueba S, Garabadzhiu AV, Martin SJ. (2018). Suppressing IL-36-driven inflammation using peptide pseudosubstrates for neutrophil proteases. https://doi.org/10.1038/s41419-018-0385-4

46 Terakawa M, Fujieda Y, Tomimori Y, Muto T, Tanaka T, Maruoka H, Nagahira K, Ogata A, Nakatsuka T, Fukuda Y. (2008). Oral chymase inhibitor SUN13834 ameliorates skin inflammation as well as pruritus in mouse model for atopic dermatitis. https://doi.org/10.1016/j.ejphar.2008.10.040

47 Badertscher K, Br&#xfffd;nnimann M, Karlen S, Braathen LRoger, Yawalkar N. (2005). Mast cell chymase is increased in chronic atopic dermatitis but not in psoriasis. https://doi.org/10.1007/s00403-005-0542-3

48 Imada T, Komorita N, Kobayashi F, Naito K, Yoshikawa T, Miyazaki M, Nakamura N, Kondo T. (2002). Development and Application of Chymase Inhibitors: Therapeutic Potential of a Specific Chymase Inhibitor in Atopic Dermatitis. https://doi.org/10.1254/jjp.90.214

49 Watanabe N, Tomimori Y, Saito K, Miura K, Wada A, Tsudzuki M, Fukuda Y. (2002). Chymase Inhibitor Improves Dermatitis in NC/Nga Mice. https://doi.org/10.1159/000064256

50 Tomimori Y, Muto T, Fukami H, Saito K, Horikawa C, Tsuruoka N, Saito M, Sugiura N, Yamashiro K, Sumida M, Kakutani S, Fukuda Y. (2002). Chymase Participates in Chronic Dermatitis by Inducing Eosinophil Infiltration. https://doi.org/10.1097/01.lab.0000018827.78602.f4

51 Kamata Y, Kimura U, Matsuda H, Tengara S, Kamo A, Umehara Y, Iizumi K, Kawasaki H, Suga Y, Ogawa H, Tominaga M, Takamori K. (2016). Relationships among plasma granzyme B level, pruritus and dermatitis in patients with atopic dermatitis. https://doi.org/10.1016/j.jdermsci.2016.09.009

52 Yawalkar N, Schmid S, Braathen LR, Pichler WJ. (2001). Perforin and granzyme B may contribute to skin inflammation in atopic dermatitis and psoriasis. https://doi.org/10.1046/j.1365-2133.2001.04222.x

53 Jung K, Pawluk MA, Lane M, Nabai L, Granville DJ. (2022). Granzyme B in epithelial barrier dysfunction and related skin diseases. https://doi.org/10.1152/ajpcell.00052.2022

54 Turner CT, Lim D, Granville DJ. (2019). Granzyme B in skin inflammation and disease. https://doi.org/10.1016/j.matbio.2017.12.005

55 Gleave A, Granville DJ. (2023). Granzyme B in Autoimmune Skin Disease. https://doi.org/10.3390/biom13020388

56 Liu Z, Shapiro SD, Zhou X, Twining SS, Senior RM, Giudice GJ, Fairley JA, Diaz LA. (2000). A critical role for neutrophil elastase in experimental bullous pemphigoid. https://doi.org/10.1172/jci3693

57 Hashimoto T, Rosen JD, Sanders KM, Yosipovitch G. (2018). Possible role of neutrophils in itch. https://doi.org/10.1097/itx.0000000000000017

58 Keith YH, Egawa G, Honda T, Kabashima K. (2023). Mast cells in type 2 skin inflammation: Maintenance and function. https://doi.org/10.1002/eji.202250359

59 Sprinzl B, Greiner G, Uyanik G, Arock M, Haferlach T, Sperr WR, Valent P, Hoermann G. (2021). Genetic Regulation of Tryptase Production and Clinical Impact: Hereditary Alpha Tryptasemia, Mastocytosis and Beyond. https://doi.org/10.3390/ijms22052458

60 Guo CJ, Mack MR, Oetjen LK, Trier AM, Council ML, Pavel AB, Guttman-Yassky E, Kim BS, Liu Q. (2020). Kallikrein 7 Promotes Atopic Dermatitis-Associated Itch Independently of Skin Inflammation. https://doi.org/10.1016/j.jid.2019.10.022

61 Zani MB, Sant'Ana AM, Tognato RC, Chagas JR, Puzer L. (2022). Human Tissue Kallikreins-Related Peptidases Are Targets for the Treatment of Skin Desquamation Diseases. https://doi.org/10.3389/fmed.2021.777619

62 Ständer S, Luger T, Kim B, Lerner E, Metz M, Adiri R, Canosa JM, Cha A, Yosipovitch G. (2024). Cutaneous Components Leading to Pruritus, Pain, and Neurosensitivity in Atopic Dermatitis: A Narrative Review. https://doi.org/10.1007/s13555-023-01081-0

63 Chavarria-Smith J, Chiu CPC, Jackman JK, Yin J, Zhang J, Hackney JA, Lin W-Y, Tyagi T, Sun Y, Tao J, Dunlap D, Morton WD, Ghodge SV, Maun HR, Li H, Hernandez-Barry H, Loyet KM, Chen E, Liu J, Tam C, Yaspan BL, Cai H, Balazs M, Arron JR, Li J, Wittwer AJ, Pappu R, Austin CD, Lee WP, Lazarus RA, Sudhamsu J, Koerber JT, Yi T. (2022). Dual antibody inhibition of KLK5 and KLK7 for Netherton syndrome and atopic dermatitis. https://doi.org/10.1126/scitranslmed.abp9159

64 Kuwae K. (2002). Epidermal expression of serine protease, neuropsin (KLK8) in normal and pathological skin samples. https://doi.org/10.1136/mp.55.4.235

65 Chen J-Q, Liang B-H, Li H-P, Mo Z-Y, Zhu H-L. (2019). Roles of Kallikrein-Related Peptidase in Epidermal Barrier Function and Related Skin Diseases. https://doi.org/10.1097/jd9.0000000000000036

66 Gouin O, Barbieux C, Leturcq F, Bonnet des Claustres M, Petrova E, Hovnanian A. (2020). Transgenic Kallikrein 14 Mice Display Major Hair Shaft Defects Associated with Desmoglein 3 and 4 Degradation, Abnormal Epidermal Differentiation, and IL-36 Signature. https://doi.org/10.1016/j.jid.2019.10.026

67 Rodriguez-Valle M, Xu T, Kurscheid S, Lew-Tabor AE. (2015). Rhipicephalus microplus serine protease inhibitor family: annotation, expression and functional characterisation assessment. https://doi.org/10.1186/s13071-014-0605-4