# Diverse viral proteases activate the NLRP1 inflammasome

**Authors**

Brian V. Tsu[1†], Christopher Beierschmitt[1†], Andrew P. Ryan[1], Rimjhim Agarwal[2], Patrick S. Mitchell[2], Matthew D. Daugherty[1*]

[†]These authors contributed equally to this work

[1]Division of Biological Sciences, University of California, San Diego, La Jolla, CA

[2]Division of Immunology & Pathogenesis, University of California, Berkeley, CA

* To whom correspondence should be addressed: mddaugherty@ucsd.edu

Keywords:

Host-virus evolution, NLRP1 inflammasome, effector-triggered immunity, pathogen-encoded proteases, picornaviruses

## ABSTRACT

The NLRP1 inflammasome is a multiprotein complex that is a potent activator of inflammation. Mouse NLRP1B can be activated through proteolytic cleavage by the bacterial Lethal Toxin (LeTx) protease, resulting in degradation of the N-terminal domains of NLRP1B and liberation of the bioactive C-terminal domain, which includes the caspase activation and recruitment domain (CARD). However, a natural pathogen-derived effector that can activate human NLRP1 remains unknown. Here, we use an evolutionary model to identify several proteases from diverse picornaviruses that cleave human NLRP1 within a rapidly evolving region of the protein, leading to host-specific and virus-specific activation of the NLRP1 inflammasome. Our work demonstrates that NLRP1 acts as a “tripwire” to recognize the enzymatic function of a wide range of viral proteases, and suggests that host mimicry of viral polyprotein cleavage sites can be an evolutionary strategy to activate a robust inflammatory immune response.

## INTRODUCTION

The ability to sense and respond to pathogens is central to the mammalian immune system. However, immune activation needs to be properly calibrated, as an overactive immune response can at times be as pathogenic as the pathogen itself. To ensure accurate discrimination of self and non-self, innate immune sensors detect broadly conserved microbial molecules such as bacterial flagellin or double-stranded RNA (Janeway, 1989). However, such microbial patterns can be found on harmless and pathogenic microbes alike. More recently, pathogen-specific activities such as toxins or effector enzymes have also been shown to be targets of innate immune recognition (Jones, Vance, & Dangl, 2016; Mitchell, Sandstrom, & Vance, 2019; Vance, Isberg, & Portnoy, 2009). Such a system for detection, termed effector-trigged immunity (ETI), has been well-established in plants (Cui, Tsuda, & Parker, 2015; Jones et al., 2016) and is emerging as an important means to allow the immune system to distinguish pathogens from harmless microbes in animals (Fischer, Naseer, Shin, & Brodsky, 2020; Jones et al., 2016).

Complicating the success of host detection systems, innate immune sensors are under constant selective pressure to adapt due to pathogen evasion or antagonism of immune detection. Such evolutionary dynamics, termed host-pathogen arms races, result from genetic conflicts where both host and pathogen are continually driven to adapt to maintain a fitness advantage. The antagonistic nature of these conflicts can be distinguished via signatures of rapid molecular evolution at the exact sites where host and pathogen interact (Daugherty & Malik, 2012; Meyerson & Sawyer, 2011; Sironi, Cagliani, Forni, & Clerici, 2015). Consistent with their role as the first line of cellular defense against incoming pathogens, innate immune sensors of both broad molecular patterns as well as specific pathogen-associated effectors have been shown to be engaged in genetic conflicts with pathogens (Cagliani et al., 2014; Chavarria-Smith, Mitchell, Ho, Daugherty, & Vance, 2016; Hancks, Hartley, Hagan, Clark, &

Elde, 2015; Tenthorey, Kofoed, Daugherty, Malik, & Vance, 2014; Tian, Pascal, & Monget, 2009).

Inflammasomes are one such group of rapidly evolving cytosolic immune sensor complexes (Broz & Dixit, 2016; Chavarria-Smith et al., 2016; Evavold & Kagan, 2019; Rathinam & Fitzgerald, 2016; Tenthorey et al., 2014; Tian et al., 2009). Upon detection of microbial molecules or pathogen-encoded activities, inflammasome-forming sensor proteins serve as a platform for the recruitment and activation of proinflammatory caspases including caspase-1 (CASP1) through either a pyrin domain (PYD) or a caspase activation and recruitment domain (CARD) (Broz & Dixit, 2016; Rathinam & Fitzgerald, 2016). Active CASP1 mediates the maturation and release of the proinflammatory cytokines interleukin (IL)-1β and IL-18 (Broz & Dixit, 2016; Rathinam, Vanaja, & Fitzgerald, 2012). CASP1 also initiates a form of cell death known as pyroptosis (Broz & Dixit, 2016; Rathinam et al., 2012). Together, these outputs provide robust defense against a wide array of eukaryotic, bacterial and viral pathogens (Broz & Dixit, 2016; Evavold & Kagan, 2019; Rathinam & Fitzgerald, 2016).

The first described inflammasome is scaffolded by the sensor protein NLRP1, a member of the nucleotide binding domain (NBD), leucine-rich repeat (LRR)-containing (NLR) superfamily (Martinon, Burns, & Tschopp, 2002; Ting et al., 2008). NLRP1 has an unusual domain architecture, containing a CARD at its C-terminus rather than the N-terminus like all other inflammasome sensor NLRs, and a function-to-find domain (FIIND), which is located between the LRRs and CARD (Ting et al., 2008). The FIIND undergoes a constitutive self-cleavage event, such that NLRP1 exists in its non-activated state as two, noncovalently associated polypeptides (D'Osualdo et al., 2011; Finger et al., 2012; Frew, Joag, & Mogridge, 2012), the N-terminal domains and the C-terminal CARD-containing fragment.

The importance of the unusual domain architecture of NLRP1 for mounting a pathogen-specific inflammasome response has been elucidated over the last several decades (Evavold & Kagan, 2019; Mitchell et al., 2019; Taabazuing, Griswold, & Bachovchin, 2020). The first hint

that NLRP1 does not detect broadly conserved microbial molecules came from the discovery that the *Bacillus anthracis* Lethal Toxin (LeTx) is required to elicit a protective inflammatory response against *B. anthracis* infection via one of the mouse NLRP1 homologs, NLRP1B (Boyden & Dietrich, 2006; Greaney et al., 2020; Moayeri et al., 2010; Terra et al., 2010). Paradoxically, inflammasome activation is the result of site-specific cleavage in the N-terminus of mouse NLRP1B by the Lethal Factor (LF) protease subunit of LeTx, indicating that protease-mediated cleavage of NLRP1 does not disable its function but instead results in its activation (Chavarria-Smith & Vance, 2013; Levinsohn et al., 2012). More recently, the mechanism by which LF-mediated proteolytic cleavage results in direct NLRP1B inflammasome activation has been detailed (Chui et al., 2019; Sandstrom et al., 2019). These studies revealed that proteolysis of mouse NLRP1B by LF results in exposure of a neo-N-terminus, which is then targeted for proteasomal degradation by a protein quality control mechanism called the “N-end rule” pathway (Chui et al., 2019; Sandstrom et al., 2019; Wickliffe, Leppla, & Moayeri, 2008; Xu et al., 2019). Since the proteasome is a processive protease, it progressively degrades the N-terminal domains of NLRP1B, but is disengaged upon arriving at the self-cleavage site within the FIIND domain. Degradation of the N-terminal domains thus releases the bioactive C-terminal CARD-containing fragment of NLRP1B from its non-covalent association with the N-terminal domains, which is sufficient to initiate downstream inflammasome activation (Chui et al., 2019; Sandstrom et al., 2019). By directly coupling NLRP1 inflammasome activation to cleavage by a pathogen-encoded protease, NLRP1 can directly sense and respond to the activity of a pathogen effector. Such a model indicates that the N-terminal domains are not required for NLRP1 activation per se, but rather serve a pathogen-sensing function. Interestingly, the N-terminal “linker” region in mouse NLRP1B that is cleaved by LF is rapidly evolving in rodents, and the analogous linker region is likewise rapidly evolving in primate species (Chavarria-Smith et al., 2016). These data suggest that selection from pathogens has

been driving diversification of this protease target region of NLRP1, possibly serving to bait diverse pathogenic proteases into cleaving NLRP1 and activating the inflammasome responses.

Consistent with the rapid evolution in NLRP1 at the site of proteolytic cleavage, LF neither cleaves nor activates human NLRP1 (Mitchell et al., 2019; Moayeri, Sastalla, & Leppla, 2012; Taabazuing et al., 2020). Despite this, human NLRP1 can also be activated by proteolysis when a tobacco etch virus (TEV) protease site is engineered into the rapidly evolving linker region of human NLRP1 that is analogous to the site of LF cleavage in mouse NLRP1B (Chavarria-Smith et al., 2016). Thus, like mouse NLRP1B, it has been predicted that human NLRP1 may serve as a "tripwire" sensor for pathogen-encoded proteases (Mitchell et al., 2019). However, direct pathogen effectors that trigger human NLRP1 have remained unidentified.

Here we investigate the hypothesis that viral proteases cleave and activate human NLRP1. We reasoned that human viruses, many of which encode proteases as necessary enzymes for their life cycle, may be triggers for NLRP1 activation. To pursue this hypothesis, we focused on viruses in the *Picornaviridae* family, which encompass a diverse set of human enteric and respiratory pathogens including coxsackieviruses, polioviruses and rhinoviruses (Zell, 2018). These viruses all translate their genome as a single polyprotein, which is then cleaved into mature proteins in at least six sites in a sequence-specific manner by a virally-encoded 3C protease, termed $3C^{pro}$ (Laitinen et al., 2016; Solomon et al., 2010; Sun, Chen, Cheng, & Wang, 2016; Zell, 2018). $3C^{pro}$ is also known to proteolytically target numerous host factors, many of which are associated with immune modulation (Croft, Walker, & Ghildyal, 2018; Huang et al., 2015; Lei et al., 2017; Mukherjee et al., 2011; Qian et al., 2017; C. Wang et al., 2019; D. Wang et al., 2012; D. Wang et al., 2014; H. Wang et al., 2015; Wen et al., 2019; Xiang et al., 2014; Xiang et al., 2016; Zaragoza et al., 2006). Because $3C^{pro}$ are evolutionarily constrained to cleave several specific polyprotein sites and host targets for replicative success, we reasoned that human NLRP1 may have evolved to sense viral infection by mimicking viral polyprotein cleavage sites, leading to NLRP1 cleavage and inflammasome activation. Using an

136 evolution-guided approach, we now show that NLRP1 is cleaved in its rapidly evolving linker
137 region by several $3C^{pro}$ from picornaviruses, resulting in inflammasome activation. We find that
138 variation in the cleavage site among primates, and even within the human population, leads to
139 changes in cleavage susceptibility and inflammasome activation. Interestingly, we also observe
140 that proteases from multiple genera of viruses cleave and human NLRP1 and mouse NLRP1B
141 at different sites, supporting a role for an evolutionary conflict between viral proteases and
142 NLRP1. Taken together, our work highlights the role of NLRP1 in sensing and responding to
143 diverse viral proteases by evolving cleavage motifs that mimic natural sites of proteolytic
144 cleavage in the viral polyprotein.

## RESULTS

### Human NLRP1 contains mimics of viral protease cleavage sites

Our hypothesis that NLRP1 can sense viral proteases is based on two prior observations. First, both human NLRP1 and mouse NLRP1B can be activated by N-terminal proteolysis (Chavarria-Smith et al., 2016). Second, the linker in primate NLRP1, which is analogous to the N-terminal disordered region of NLRP1B that is cleaved by LF protease, has undergone recurrent positive selection (Chavarria-Smith et al., 2016), or an excess of non-synonymous amino acid substitutions over what would be expected by neutral evolution (Kimura, 1983). We reasoned that this rapid protein sequence evolution may reflect a history of pathogen-driven selection, wherein primate NLRP1 has evolved to sense pathogen-encoded proteases such as those encoded by picornaviruses. To test this hypothesis, we first generated a predictive model for $3C^{pro}$ cleavage site specificity. We chose to focus on the enterovirus genus of picornaviruses, as there is a deep and diverse collection of publicly available viral sequences within this genus due to their importance as human pathogens including coxsackieviruses, polioviruses, enterovirus D68 and rhinoviruses (Blom, Hansen, Blaas, & Brunak, 1996; Pickett et al., 2012). We first compiled complete enterovirus polyprotein sequences from the Viral Pathogen Resource (ViPR) database (Pickett et al., 2012) and extracted and concatenated sequences for each of the cleavage sites within the polyproteins (Figure 1A and 1B, Supplementary files 1 and 2). After removing redundant sequences, we used the MEME Suite (Bailey et al., 2009) to create the following $3C^{pro}$ cleavage motif: [A/Φ]XXQGXXX (where Φ denotes a hydrophobic residue and X denotes any amino acid), which is broadly consistent with previous studies (Blom et al., 1996; Fan et al., 2020; Jagdeo et al., 2018; O'Donoghue et al., 2012) (Figure 1C).

We next optimized our $3C^{pro}$ cleavage site motif prediction by querying against predicted viral polyprotein and experimentally validated host cleavage sites (Laitinen et al., 2016),

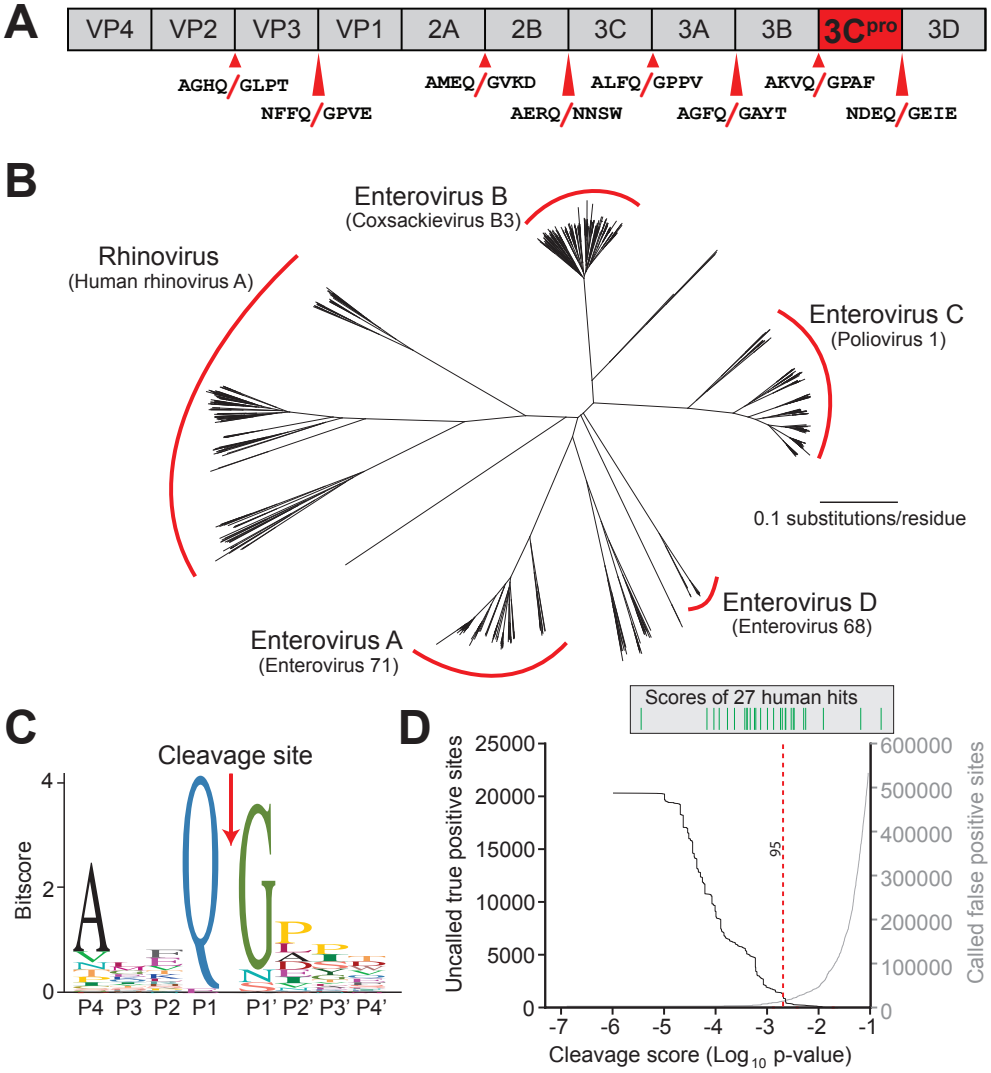


**Figure 1. Conserved polyprotein cleavage sites across enteroviruses inform substrate specificity of the enteroviral 3C$^{pro}$.** (A) Schematic of 3C$^{pro}$ cleavage sites (red arrows) within the polyprotein of coxsackievirus B3 Nancy (CVB3), a model enterovirus. Shown are the eight amino acids flanking each cleavage site within the polyprotein. (B) Phylogenetic tree of 796 enteroviral polyprotein coding sequences depicting the major clades of enteroviruses sampled in this study with representative viruses from each clade in parentheses (Supplementary file 2). (C) Eight amino acid polyprotein cleavage motif for enteroviruses (labeled as positions P4 to P4') generated from the 796 enteroviral polyprotein sequences in (B) using the MEME Suite (Supplementary file 2). (D) Training set data used to determine the motif search threshold for FIMO (Supplementary files 1, 3 and 4). The X-axis represents a $\log_{10}$ of the p-value reported by FIMO as an indicator for the strength of the cleavage motif hit (cleavage score). (Left) The Y-axis depicts the number of uncalled true positives, or motif hits that overlap with the initial set of 8mer polyprotein cleavage sites used to generate the motif, in the training set of enteroviral polyprotein sequences (black). (Right) The Y-axis depicts the number of called false positive sites, or any motif hits found in the polyprotein that are not known to be cleaved by 3C$^{pro}$, in the training set of enteroviral polyprotein sequences (gray). (Above) Each line depicts a single, experimentally validated case of enteroviral 3C$^{pro}$ cleavage site within a human protein as reported in Laitinen et al, 2016 and is ordered along the x-axis by its resulting cleavage score. A vertical dotted line is used to represent the decided threshold that captures 95% of true positive hits and 16 out of 27 reported human hits (Figure 1 – figure supplement 1).

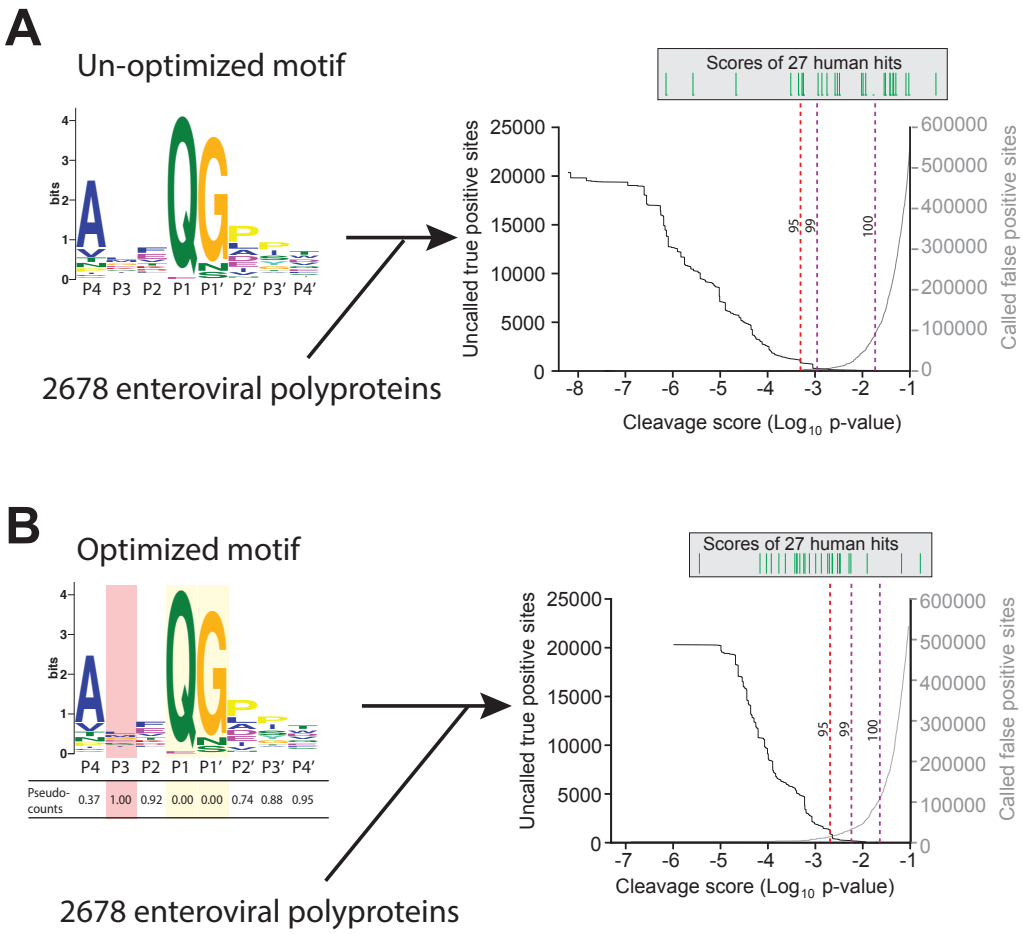


**Figure 1 – figure supplement 1.** Motif optimization enhances capture of known human targets of enteroviral 3C$^{pro}$. (A) As described in Figure 1B and 1C and Materials and Methods, the 8mer (P4-P4') 3C$^{pro}$ polyprotein cleavage motif was initially generated from unique, concatenated 8mer cleavage sites across 796 enteroviral polyprotein sequences. To assess the capture capability of the motif on both virus and host targets, the motif was then used to conduct a low threshold (p-value = 0.1) FIMO (MEME Suite) search across training set of 2678 nonredundant enteroviral polyproteins from ViPR and 27 experimentally validated human targets of 3C$^{pro}$ (Laitinen et al., 2016). In the graph, the X-axis represents a $\log_{10}$ of the p-value reported by FIMO as an indicator for the strength of the cleavage motif hit, or cleavage score. The left Y-axis depicts the number of uncalled true positives, or motif hits within the enteroviral polyprotein training set that overlap with the initial set of 8mer polyprotein cleavage sites used to generate the motif (black). The right Y-axis depicts the number of called false positive sites, or any motif hits that are not true positives, in the training set of enteroviral polyprotein sequences (gray). (Above) Each line depicts a single, experimentally validated case of enteroviral 3C$^{pro}$ cleavage site within a human protein as reported in Laitinen et al, 2016 and is ordered along the x-axis by its corresponding cleavage score. Vertical dotted lines are used to represent the decided thresholds for comparison of capture capability. Capture of human targets at 95%, 99%, or 100% capture of true positives in the polyprotein dataset corresponds to capture of 4, 7, and 16 human hits. (B) Pseudo-counts to the position-specific scoring matrix of the motif shown in (A) were adjusted by total information content where the two most information-dense positions P1 and P1' are assigned pseudocount = 0 and the least information-dense position P3 pseudocount = 1, and the remaining positions are assigned a pseudocount value relative to the most information-dense position P1. This optimized motif is then used to FIMO search against the same training set as described in (A). Capture of human targets at 95%, 99%, or 100% capture of true positives in the polyprotein dataset corresponds to capture of 16, 23, and 24 human hits.

allowing us to set thresholds for predicting new cleavage sites (Supplementary files 3 and 4). Due to the low information content of the polyprotein motif (Figure 1C), such predictions are necessarily a compromise between stringency and capturing the most known cleavage sites. In particular, we wished to make sure that the model was able to capture a majority of experimentally validated human hits (compiled in (Laitinen et al., 2016)) in addition to the known sites of polyprotein cleavage ("true positives"), while minimizing the prediction of sites outside of known polyprotein cleavage sites ("false positives"). By adjusting the model to allow greater flexibility for amino acids not sampled in the viral polyprotein (see Methods and Figure 1 – figure supplement 1 and Supplementary file 4), we were able to capture 95% of known viral sites and the majority of the known human hits, while limiting the number of false negative hits within the viral polyprotein (Figure 1D).

**The coxsackievirus B3 3C$^{pro}$ cleaves and activates human NLRP1 at a predicted site within the linker region.**

We next used our refined model to conduct a motif search for 3C$^{pro}$ cleavage sites in NLRP1 using Find Individual Motif Occurrences (FIMO) (Grant, Bailey, & Noble, 2011). We identified three occurrences of the motif across the full-length human NLRP1 protein (Figure 2A). Of these sites, one in particular, 127-GCTQGSER-134, fell within the previously described rapidly-evolving linker (Chavarria-Smith et al., 2016) and demonstrates the lowest percent conservation across mammalian species at each of the predicted P4-P4' positions (Figure 2B).

To assess if human NLRP1 is cleaved by enteroviral 3C$^{pro}$, we co-expressed a N-terminal mCherry-tagged wild-type (WT) human NLRP1 with the 3C$^{pro}$ from the model enterovirus, coxsackievirus B3 (CVB3) in HEK293T cells (Figure 2C). The mCherry tag stabilizes and allows visualization of putative N-terminal cleavage products, similar to prior studies (Chavarria-Smith et al., 2016). We observed that the WT but not catalytically inactive (C147A) CVB3 3C$^{pro}$ cleaved NLRP1, resulting in a cleavage product with a molecular weight

consistent with our predicted $3C^{pro}$ cleavage at the predicted 127-GCTQGSER-134 site (44 kDa) (Figure 2D). Based on the presence of a single cleavage product, we assume that the other predicted sites are either poor substrates for $3C^{pro}$ or less accessible to the protease as would be predicted from their NetSurfP-reported (Klausen et al., 2019) coil probability within structured domains of the protein (Figure 2A and Figure 2 – figure supplement 1). To determine if the cleavage occurs between residues 130 and 131, we mutated the P1' glycine to a proline (G131P), which abolished $3C^{pro}$ cleavage of NLRP1 (Figure 2D). CVB3 $3C^{pro}$ cleavage of NLRP1 resulted in a similarly intense cleavage product when compared to the previously described system in which a TEV protease site was introduced into the linker region of NLRP1 (Chavarria-Smith et al., 2016) (Figure 2D). Taken together, these results indicate that cleavage of WT NLRP1 by a protease from a natural human pathogen is robust and specific.

During a viral infection, $3C^{pro}$ is generated in the host cell cytoplasm after translation of the viral mRNA to the polyprotein and subsequent processing of the viral polyprotein into constituent pieces (Laitinen et al., 2016). To confirm that $3C^{pro}$ generated during a viral infection is able to cleave NLRP1, we virally infected cells expressing either WT NLRP1 or the uncleavable (G131P) mutant. We observed accumulation of the expected cleavage product beginning at six hours post-infection when we infected cells expressing WT NLRP1 and no cleavage product when we infected cells expressing the 131P mutant (Figure 2E). These results validate that CVB3 infection can result in rapid and specific cleavage of human NLRP1.

Previous results with a TEV-cleavable human NLRP1 showed that cleavage by TEV protease was sufficient to activate the human NLRP1 inflammasome in a reconstituted inflammasome assay (Chavarria-Smith et al., 2016). Using the same assay, in which plasmids-encoding human NLRP1, CASP1, ASC and IL-1β are transfected into HEK293T cells, we tested if the CVB3 $3C^{pro}$ activates the NLRP1 inflammasome. We observed that the CVB3 $3C^{pro}$ results in robust NLRP1 inflammasome activation, as measured by CASP1-dependent processing of pro-IL-1β to the active p17 form (Figure 2F). To confirm that $3C^{pro}$-induced inflammasome

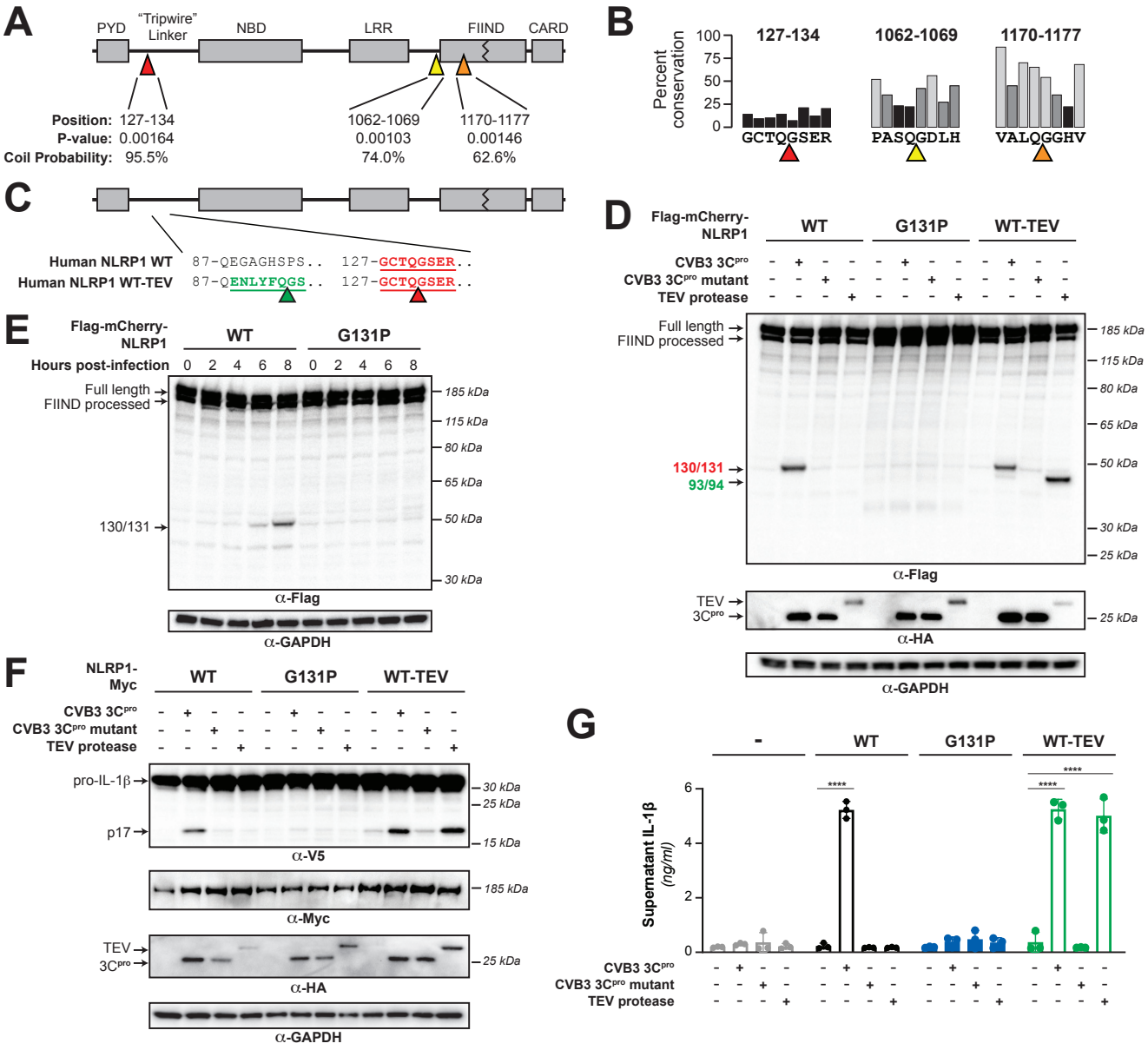


**Figure 2. Enterovirus 3C$^{pro}$ cleaves human NLRP1 at the predicted site of mimicry and promotes pro-inflammatory cytokine release.** (A) Schematic of the domain structure of NLRP1, with predicted cleavage sites (triangles). FIMO-reported p-values and average NetsurfP-reported coil probabilities (Figure 2 – figure supplement 1) are described at the predicted sites. (B) Percent conservation across 100 mammalian species at each position of each predicted 8mer cleavage site within human NLRP1. (C) Schematic of the human NLRP1 sequence used to assess enteroviral cleavage and activation. The predicted enteroviral cleavage site found in the linker region (127-GCTQGSER-134) is shown in red. Human NLRP1 WT-TEV contains an engineered TEV cleavage site between residues 93 and 94 (underlined green) in human NLRP1 WT. (D) Immunoblot depicting human NLRP1 cleavage by CVB3 3C$^{pro}$ and TEV protease. HEK293T cells were co-transfected using 100ng of the indicated Flag-tagged mCherry-NLRP1 fusion plasmid constructs with 250ng of the indicated protease construct and immunoblotted with the indicated antibodies. (E) Immunoblot depicting human NLRP1 cleavage at the indicated timepoints after infection with 250,000 PFU (MOI = ~1) CVB3. HEK293T cells were transfected using 100ng of either WT NLRP1 or NLRP1 G131P and infected 24-30 hours later. All samples were harvested 32 hours post-transfection and immunoblotted with the indicated antibodies (F) Immunoblot depicting human NLRP1 activation (maturation of IL-1β) by CVB3 3C$^{pro}$ and TEV protease. HEK293T cells were co-transfected using 100ng of the indicated protease, 50ng V5-IL-1β,

100ng CASP1, 5ng ASC, and 4ng of the indicated Myc-tagged NLRP1, and immunoblotted with the indicated antibodies. Appearance of the mature p17 band of IL-1β indicates successful assembly of the NLRP1 inflammasome and activation of CASP1. (G) Bioactive IL-1β in the culture supernatant was measured using HEK-Blue IL-1β reporter cells, which express secreted embryonic alkaline phosphatase (SEAP) in response to extracellular IL-1β. Supernatant from cells transfected as in (E) was added to HEK-Blue IL-1β reporter cells and SEAP levels in the culture supernatant from HEK-Blue IL-1β reporter cells were quantified by the QUANTI-Blue colorimetric substrate. Transfections were performed in triplicate and compared to the standard curve generated from concurrent treatment of HEK-Blue IL-1β reporter cells with purified human IL-1β (Figure 2 – figure supplement 2). Data were analyzed using two-way ANOVA with Sidak's post-test. **** = $p < 0.0001$.

| Class: Buried (B), Exposed (E) | Amino Acid Position | Amino Acid | Relative Surface Accessibility | Absolute Surface Accessibility | Probability For Alpha-Helix | Probability For Beta-strand | Coil Probability |
|---|---|---|---|---|---|---|---|
| E | 127 | G | 0.724 | 56.997 | 0.003 | 0.002 | 0.994 |
| E | 128 | C | 0.450 | 63.190 | 0.004 | 0.006 | 0.990 |
| E | 129 | T | 0.607 | 84.168 | 0.004 | 0.014 | 0.982 |
| E | 130 | Q | 0.597 | 106.589 | 0.004 | 0.022 | 0.974 |
| E | 131 | G | 0.534 | 42.063 | 0.012 | 0.016 | 0.972 |
| E | 132 | S | 0.601 | 70.429 | 0.018 | 0.023 | 0.960 |
| E | 133 | E | 0.557 | 97.244 | 0.025 | 0.068 | 0.908 |
| E | 134 | R | 0.456 | 104.412 | 0.016 | 0.127 | 0.857 |
| | | | | | | | |
| E | 1062 | P | 0.311 | 44.067 | 0.353 | 0.016 | 0.631 |
| E | 1063 | A | 0.435 | 47.991 | 0.389 | 0.029 | 0.582 |
| E | 1064 | S | 0.307 | 35.976 | 0.372 | 0.038 | 0.590 |
| E | 1065 | Q | 0.406 | 72.423 | 0.204 | 0.050 | 0.746 |
| E | 1066 | G | 0.384 | 30.211 | 0.136 | 0.068 | 0.797 |
| E | 1067 | D | 0.321 | 46.271 | 0.103 | 0.079 | 0.818 |
| E | 1068 | L | 0.321 | 58.772 | 0.058 | 0.055 | 0.887 |
| E | 1069 | H | 0.459 | 83.520 | 0.047 | 0.081 | 0.872 |
| | | | | | | | |
| B | 1170 | V | 0.085 | 13.065 | 0.087 | 0.811 | 0.102 |
| B | 1171 | A | 0.062 | 6.837 | 0.085 | 0.769 | 0.146 |
| B | 1172 | L | 0.165 | 30.146 | 0.087 | 0.399 | 0.514 |
| E | 1173 | Q | 0.519 | 92.645 | 0.078 | 0.213 | 0.709 |
| E | 1174 | G | 0.576 | 45.362 | 0.051 | 0.051 | 0.898 |
| E | 1175 | G | 0.490 | 38.555 | 0.033 | 0.028 | 0.938 |
| E | 1176 | H | 0.618 | 112.433 | 0.011 | 0.108 | 0.881 |
| B | 1177 | V | 0.144 | 22.068 | 0.008 | 0.171 | 0.821 |

**Figure 2 – figure supplement 1.**
Tabular output of NetSurfP structural predictions for human NLRP1 describing the predicted class (buried or exposed), relative surface accessibility, absolute surface accessibility, probability for alpha-helix, probability for beta-strand, and coil probability for the amino acid positions within each predicted enteroviral $3C^{pro}$ cleavage site.

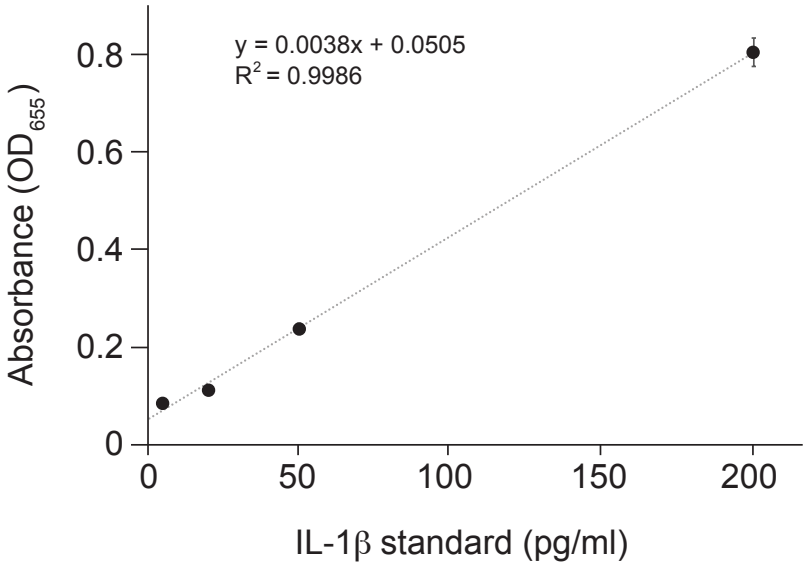


**Figure 2 – figure supplement 2.**
Standard curve for Figure 2G. Purified human IL-1β was added in duplicate to the indicated final concentration to HEK-Blue IL-1β reporter cells and SEAP activity was measured by increased absorbance at $OD_{655}$. The indicated linear fit was used to calculate absolute concentrations of bioactive IL-1β from culture supernatants shown in Figure 2F. Note that supernatants from inflammasome-transfected cells was diluted 10-fold before addition to HEK-Blue IL-1β reporter cells to ensure that levels fell within the linear range of the indicated standard curve. Standard curves were generated in an identical manner for each panel of HEK-Blue data shown.

223 activation resulted in release of bioactive IL-1β from cells, we measured active IL-1β levels in
224 the culture supernatant using cells engineered to express a reporter gene in response to
225 soluble, active IL-1β. When compared to a standard curve (Figure 2 – figure supplement 2), we
226 found that $3C^{pro}$ treatment resulted in release of >4 ng/ml of active IL-1β into the culture
227 supernatant (Figure 2G). Importantly, in both assays, $3C^{pro}$-induced inflammasome activation
228 was comparable to TEV-induced activation and was ablated when position 131 was mutated,
229 validating that CVB3 $3C^{pro}$ cleavage at a single site is both necessary and sufficient to activate
230 NLRP1 (Figure 2F and 2G).
231
232 **NLRP1 diversification across primates and within humans confers host differences in**
233 **susceptibility to viral $3C^{pro}$ cleavage and inflammasome activation**
234 Our evolutionary model in which NLRP1 is evolving in conflict with $3C^{pro}$ suggests that
235 changes in the NLRP1 linker region, both among primates and within the human population
236 (Figure 3A), would confer host-specific differences to NLRP1 cleavage and inflammasome
237 activation. To test this hypothesis, we aligned the linker regions from NLRP1 from diverse
238 mammals and human population sampling and compared the sequences around the site of
239 CVB3 $3C^{pro}$ cleavage (Figure 3B and 3C and Figure 3 – figure supplement 1). We noted that
240 while a majority of primate NLRP1s are predicted to be cleaved similarly to the human ortholog,
241 several primate proteins would be predicted to not be cleaved by enteroviral $3C^{pro}$ as a result of
242 changes to either the P4, P1 or P1' residues. To confirm these predictions, we made the human
243 NLRP1 mutants G127E or G131R, which reflect the Old World monkey or marmoset residues at
244 each position, respectively. As predicted, both primate NLRP1 variants prevented $3C^{pro}$
245 cleavage of NLRP1 (Figure 3D). These results indicate that multiple viral $3C^{pro}$ activate host
246 NLRP1 in a host specific manner and suggest that single changes within a short linear motif can
247 substantially alter cleavage susceptibility and inflammasome activation.

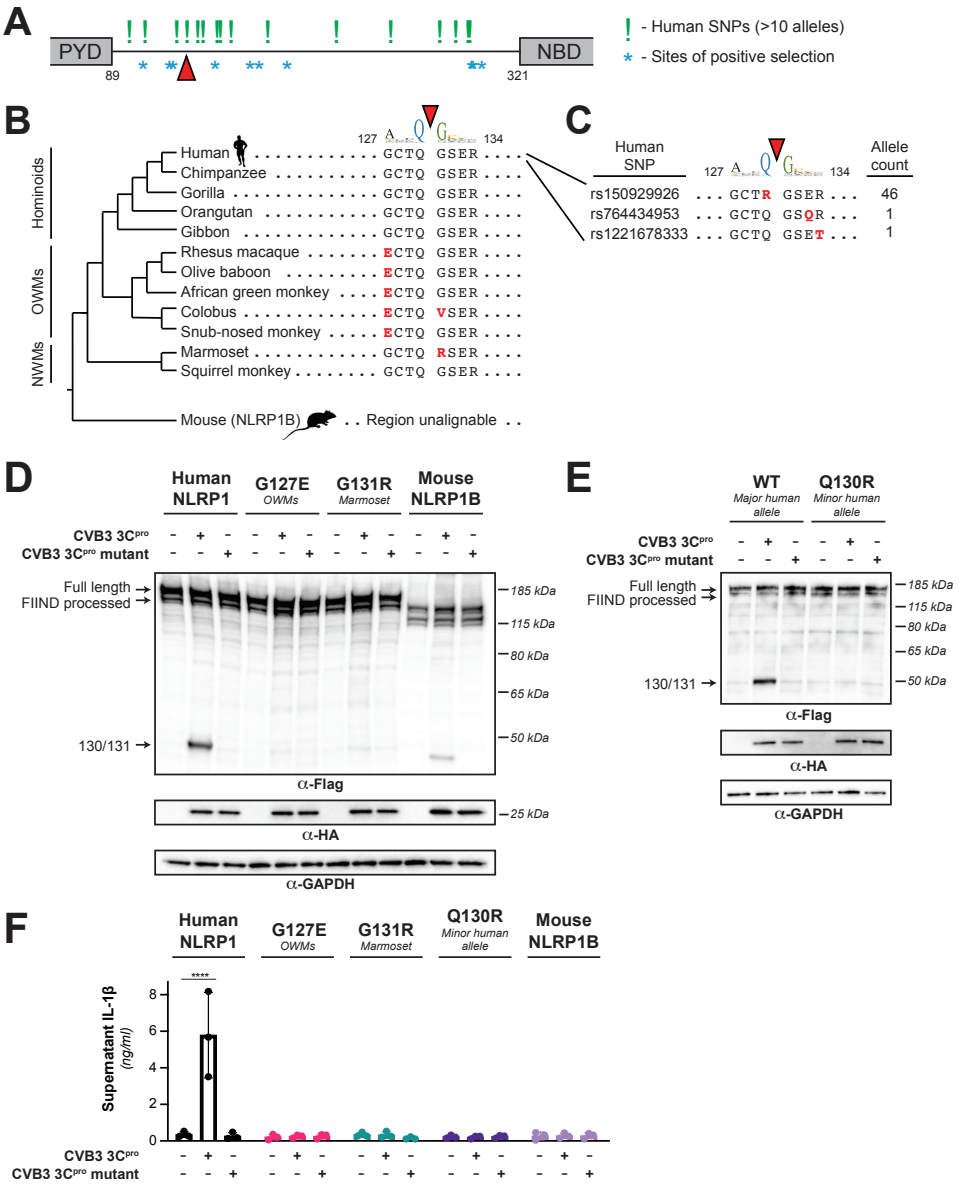


**Figure 3. Naturally occurring cleavage site variants alter NLRP1 susceptibility to enteroviral 3C$^{pro}$.** (A) Schematic of sites found to be evolving under positive selection (marked as *, from (Chavarria-Smith & Vance, 2013)) and human SNPs with at least 10 reported instances in the Genome Aggregation Database (GnomAD, (Karczewski et al., 2020)) (marked as !) within the linker region between the pyrin domain (PYD) and nucleotide binding domain (NBD) of NLRP1. The enteroviral 3C$^{pro}$ cleavage site between position 130 and 131 is indicated by a red triangle. (B) Phylogenetic tree depicting the enteroviral 3C$^{pro}$ cleavage site (red triangle) within NLRP1 across three clades of primates – hominoids, Old World monkeys (OWMs), and New World monkeys (NWMs). Mouse NLRP1B lacks any sequence that is alignable to this region of primate NLRP1 (see also Figure 3 – figure supplements 1 & 2). Amino acid differences to the human NLRP1 reference sequence are highlighted in red. Above the alignment is the enterovirus 3C$^{pro}$ sequence logo shown in Figure 1. (C) GnomAD-derived allele counts of each missense human SNP (by reference SNP #) within the 8mer of the determined enteroviral 3C$^{pro}$ cleavage site. (D-E) Immunoblot depicting CVB3 3C$^{pro}$ cleavage susceptibility of the indicated 8mer site variants introduced into human NLRP1 or full-length wild-type mouse NLRP1B (D) or the cleavage susceptibility of human NLRP1 Q130R, a naturally occurring human population variant (E). (F) Release of bioactive IL-1β into the culture supernatant as measured using HEK-Blue IL-1β reporter cells as in Figure 2F. Data were analyzed using two-way ANOVA with Sidak's post-test. **** = $p < 0.0001$.

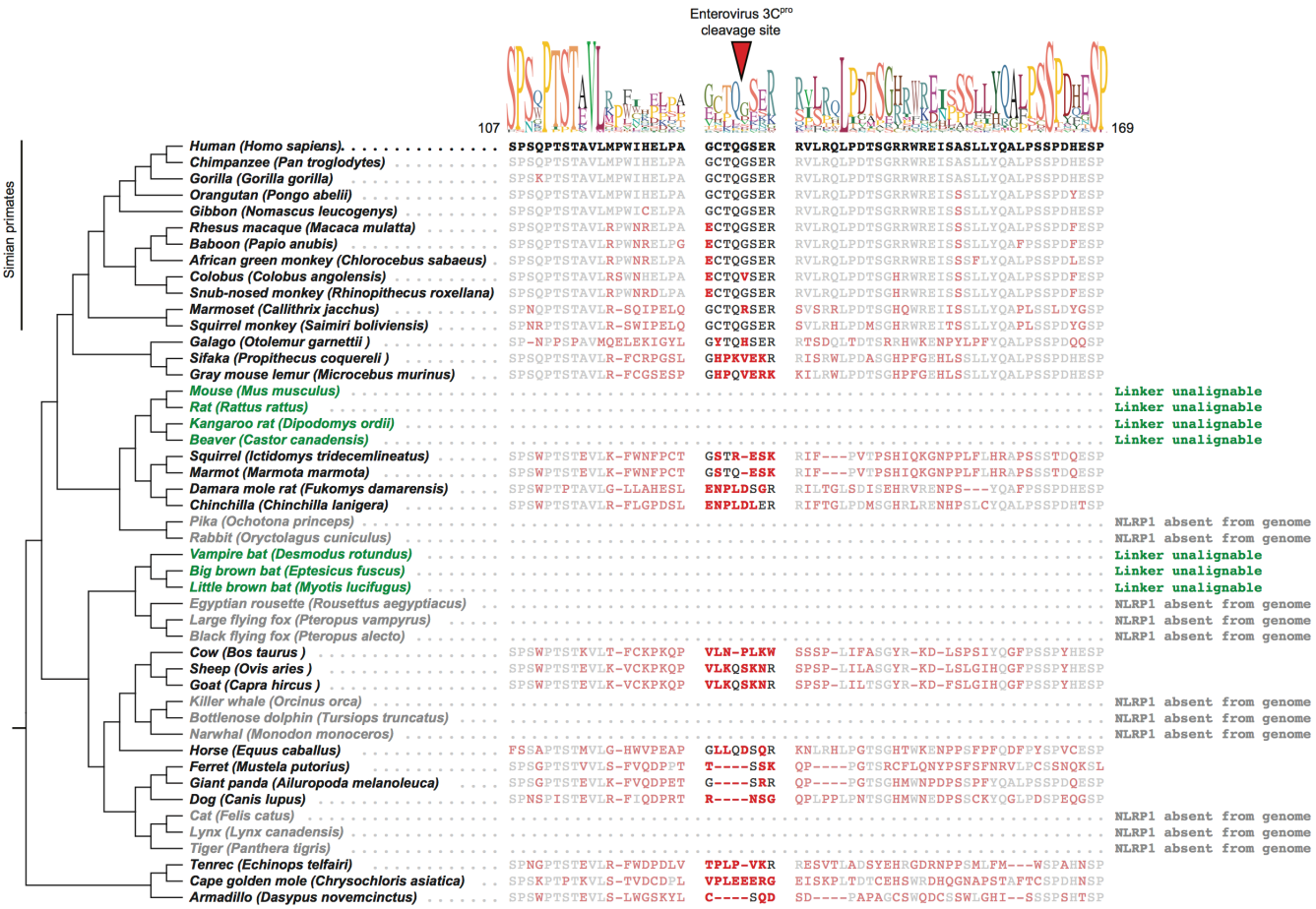


**Figure 3 – figure supplement 1.**

Mammalian NLRP1 phylogenomics and alignment of linker region. The indicated mammalian NLRP1 sequences were aligned and the region corresponding to residues 107-169 from human NLRP1 were extracted, which is anchored on both ends by well conserved proline and serine rich motifs. A consensus sequence generated from alignable sequences in this region is shown above the human sequence. The position of the CVB3 3C$^{pro}$ cleavage site in human NLRP1 is shown, flanked by four amino acids on both sides (P4->P4'). In other mammals, residues that differ from the human sequence are shown in red. Within the aligned region that corresponds to the CVB3 3C$^{pro}$ cleavage site, only simian primates have P4, P1 and P1' residues that would allow cleavage. The only other species that have a plausible cleavage site in this position are sheep and goats (P4 = Val, P1 = Gln, P1' = Ser), although those residues appear to have evolved independently at those positions. Two clades of species (the "mouse-related" clade of rodents and the microbat clade, marked as green) have NLRP1 protein sequences with N-terminal linkers that are unalignable to human NLRP1 in this region. Four additional clades (lagomorphs, megabats, cetaceans and felines, marked as grey) lack the NLRP1 gene altogether.

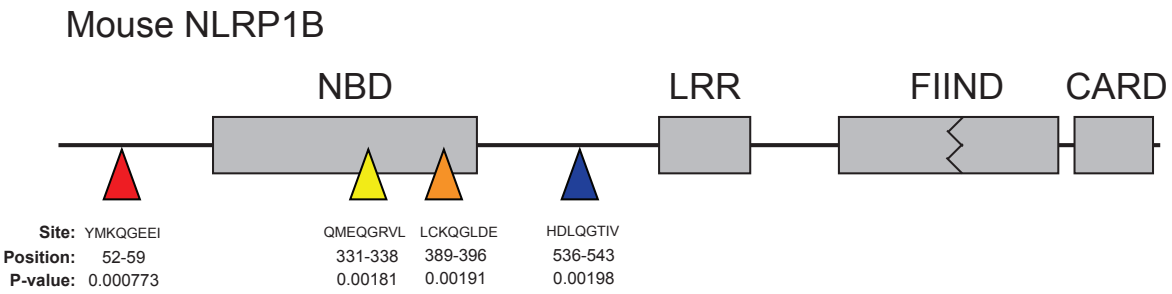


**Figure 3 – figure supplement 2.**
Schematic of the domain structure of mouse NLRP1B, with predicted cleavage sites shown as colored triangles. FIMO-reported p-values are described at the predicted sites. Note that mouse NLRP1 genes, like those of many rodents, lack the N-terminal pyrin domain found in human NLRP1.

We further observed that this cleavage site is largely absent in non-primate species (Figure 3 – figure supplement 1), suggesting that a $3C^{pro}$ cleavage site mimic emerged in simian primates 30-40 million years ago. While many other mammalian species have a region that is alignable to the primate linker, we noted that this region is unalignable to any sequence in the linker region of NLRP1 proteins from rodents or bats (Figure 3B and Figure 3 – figure supplement 1). Despite this, we found that there was weak cleavage of mouse NLRP1B at a site closer to the N-terminus than the 127-GCTQGSER-134 site found in human NLRP1 (Figure 3D), suggesting that an independent cleavage site could have arisen elsewhere in mouse NLRP1B. Indeed, our best prediction for a cleavage site in mouse NLRP1B is between residues 55 and 56 (52-YMKQGEEI-59, Figure 3 – figure supplement 2). These data suggest that NLRP1 in other mammals may have convergently evolved cleavage sites in the linker region despite not having a cleavable sequence in the precise position that human NLRP1 is cleaved.

Differential host susceptibility to NLRP1 cleavage and activation extends to the human population level. Using GnomAD (Karczewski et al., 2020), we sampled the alternative alleles within the direct cleavage site (Figure 3C). While this region does not appear to be highly polymorphic in humans, we note that one alternative allele (rs150929926) results in a Q130R mutation and is present in >1 in every 1000 African alleles sampled. Introducing this mutation into NLRP1, we find the Q130R mutation eliminates NLRP1 cleavage susceptibility to CVB3 $3C^{pro}$ (Figures 3E). In the case of primate and human diversity alleles at the site of $3C^{pro}$ cleavage, we also find that loss of cleavage susceptibility results in a loss of inflammasome activation in response to $3C^{pro}$ (Figure 3F), supporting the aforementioned notion that single changes in the linker region can have drastic impacts on the ability of different hosts to respond to the presence of cytoplasmic $3C^{pro}$.

**$3C^{pro}$ from diverse picornaviruses cleave and activate human NLRP1.**

Our evolutionary model predicted that NLRP1 would be cleaved by a broad range of $3C^{pro}$ from viruses in the enterovirus genus (Figure 1B). To test this hypothesis, we cloned $3C^{pro}$ from representative viruses from four additional major species of human enteroviruses: enterovirus 71 (EV71, species: *Enterovirus A*), poliovirus 1 (PV1, species: *Enterovirus C*), enterovirus D68 (EV68, species: *Enterovirus D*), human rhinovirus A (HRVA, species: *Rhinovirus A*), in order to compare them to the $3C^{pro}$ from CVB3 (species: *Enterovirus B*) (Figure 4A). Confirming our computational prediction, and despite <50% amino acid identity between some of these proteases (Figure 4 – figure supplement 1 and 2) we found that every tested member of enterovirus $3C^{pro}$ was able to cleave NLRP1 between residues 130 and 131 (Figure 4B). Moreover, expression of every tested enterovirus $3C^{pro}$ resulted in activation of the inflammasome in a manner that was dependent on cleavage at the 127-GCTQGSER-134 site (Figure 4C).

Enteroviruses are only one genus within the broad *Picornaviridae* family of viruses. We next asked if viruses in other *Picornaviridae* genera that infect humans are also able to cleave and activate human NLRP1. We were unable to generate a robust sequence motif for every genera of picornavirus due to lower depth of publicly available sequences. Instead, we cloned a $3C^{pro}$ from a representative of every genus of picornavirus that are known to infect humans: encephalomyocarditis virus (EMCV, genus: *Cardiovirus*), parechovirus A virus (ParA, genus: *Parechovirus*), Aichi virus (Aichi, genus: *Kobuvirus*), hepatitis A virus (HepA, genus: *Hepatovirus*), salivirus A virus (SaliA, genus: *Salivirus*), and rosavirus A2 (Rosa2, genus: *Rosavirus*). Each of these viral proteases is <20% identical to CVB3 $3C^{pro}$ (Figure 4 – figure supplement 2). Despite this, the sequence motif built from cleavage sites within the polyprotein of these individual viruses is broadly consistent with the motif seen in enteroviruses (Figure 4A), reflective of the strong evolutionary constraint on evolution of the sequence specificity of these proteases. Interestingly, we found that there was substantial variation in NLRP1 cleavage sites across these diverse $3C^{pro}$ even though most picornavirus proteases cleaved human NLRP1 to

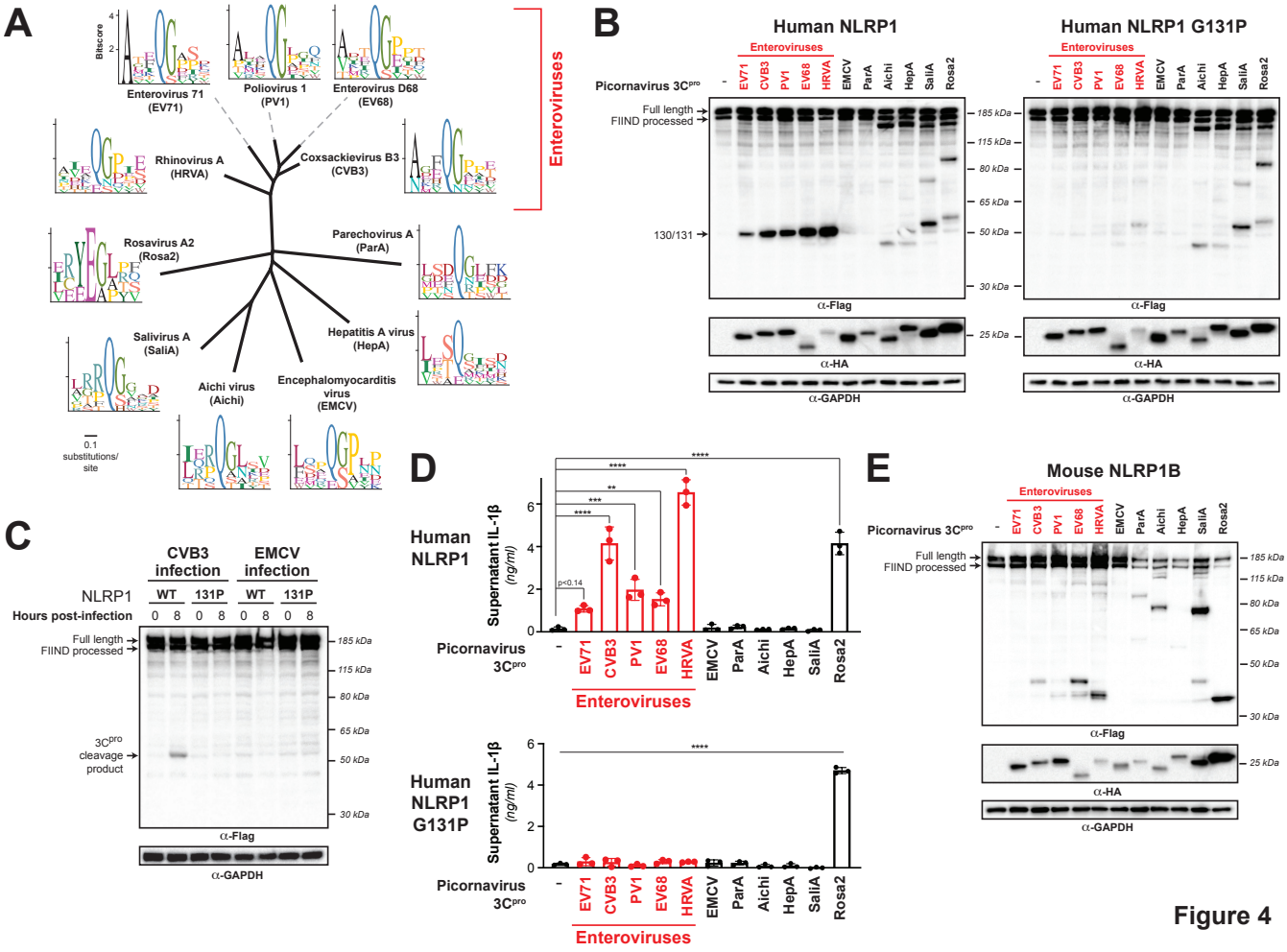


Figure 4

**Figure 4. Diverse picornavirus 3C$^{pros}$ cleave and activate NLRP1 at independently evolved sites.** (A) Phylogenetic tree of 3C$^{pro}$ protein sequences for the indicated picornaviruses (Figure 4 – figure supplement 1 and 2). Shown next to the virus name is the sequence motif generated from the known sites of 3C$^{pro}$ polyprotein cleavage in that specific virus. (B) Immunoblot depicting human NLRP1 cleavage by the indicated picornaviral 3C$^{pro}$. Abbreviations are as in (A). Assays were performed as in Figure 2D. (left) Cleavage assays against WT NLRP1. (right) Human NLRP1 G131P mutant used in Figure 2. (C) Immunoblot depicting human NLRP1 cleavage at the indicated timepoints after infection with 250,000 PFU (MOI = ~1) CVB3 or EMCV. HEK293T cells were transfected using 100ng of either WT NLRP1 or NLRP1 G131P and infected 24-30 hours later. All samples were harvested 32 hours post-transfection and immunoblotted with the indicated antibodies (D) Release of bioactive IL-1β into the culture supernatant as measured using HEK-Blue IL-1β reporter cells as in Figure 2G. Data were analyzed using one-way ANOVA with Tukey's post-test. ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. (E) Immunoblot depicting mouse NLRP1B cleavage by numerous picornaviral 3C$^{pro}$. Assays were performed as in Figure 2D.

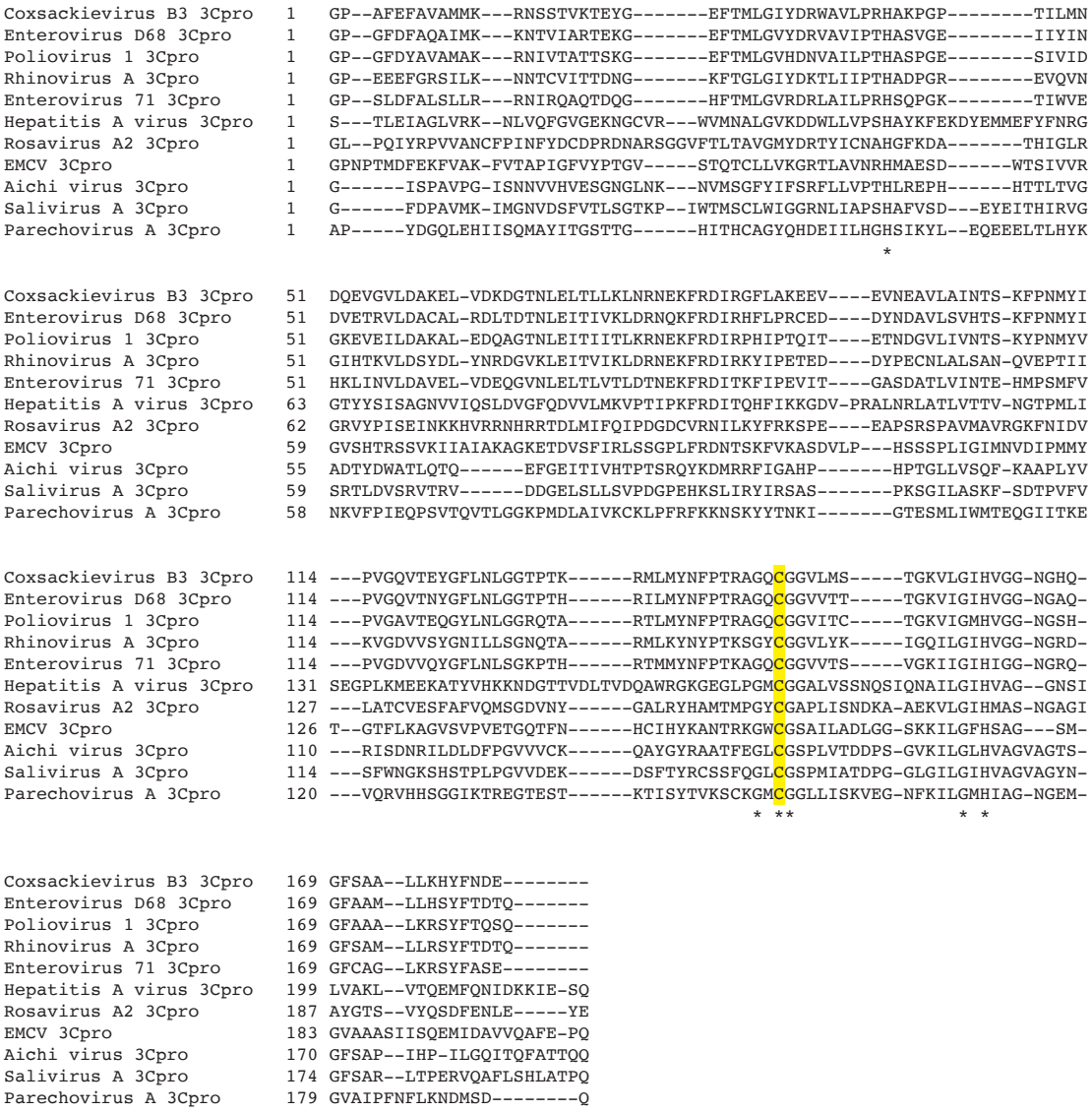


**Figure 4 – figure supplement 1.**
Alignment of $3C^{pros}$ used in this study. Sequences were aligned using MAFFT (Katoh & Standley, 2013) and used to generate the phylogenetic tree shown in Figure 4A. Asterisks indicate residues 100% conserved in all sequences. The position of the catalytic cysteine, analogous to C147 in CVB3 $3C^{pro}$, is highlighted in yellow.

**Amino acid identity between 3Cpros (%)**

| | *CVB3* | *EV68* | *PV1* | *HRVA* | *EV71* | *HepA* | *Rosa2* | *EMCV* | *Aichi* | *SaliA* | *ParA* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Coxsackievirus B3 3Cpro (CVB3)*** | | 67.0 | 61.0 | 47.3 | 54.9 | 19.2 | 18.5 | 17.8 | 19.0 | 19.1 | 16.5 |
| ***Enterovirus D68 3Cpro (EV68)*** | 67.0 | | 65.6 | 49.2 | 53.3 | 17.7 | 16.4 | 16.3 | 18.9 | 16.0 | 13.9 |
| ***Poliovirus 1 3Cpro (PV1)*** | 61.0 | 65.6 | | 43.7 | 54.9 | 17.2 | 17.4 | 19.7 | 18.4 | 14.5 | 15.4 |
| ***Rhinovirus A 3Cpro (HRVA)*** | 47.3 | 49.2 | 43.7 | | 45.6 | 17.2 | 18.9 | 16.3 | 19.9 | 20.0 | 15.4 |
| ***Enterovirus 71 3Cpro (EV71)*** | 54.9 | 53.3 | 54.9 | 45.6 | | 15.4 | 17.0 | 16.8 | 16.4 | 16.6 | 18.0 |
| ***Hepatitis A virus 3Cpro (HepA)*** | 19.2 | 17.7 | 17.2 | 17.2 | 15.4 | | 18.1 | 20.3 | 17.9 | 17.0 | 15.5 |
| ***Rosavirus A2 3Cpro (Rosa2)*** | 18.5 | 16.4 | 17.4 | 18.9 | 17.0 | 18.1 | | 17.1 | 14.3 | 16.9 | 15.7 |
| ***EMCV 3Cpro (EMCV)*** | 17.8 | 16.3 | 19.7 | 16.3 | 16.8 | 20.3 | 17.1 | | 16.1 | 23.7 | 16.0 |
| ***Aichi virus 3Cpro (Aichi)*** | 19.0 | 18.9 | 18.4 | 19.9 | 16.4 | 17.9 | 14.3 | 16.1 | | 33.8 | 13.1 |
| ***Salivirus A 3Cpro (SaliA)*** | 19.1 | 16.0 | 14.5 | 20.0 | 16.6 | 17.0 | 16.9 | 23.7 | 33.8 | | 13.5 |
| ***Parechovirus A 3Cpro (ParA)*** | 16.5 | 13.9 | 15.4 | 15.4 | 18.0 | 15.5 | 15.7 | 16.0 | 13.1 | 13.5 | |

**Figure 4 – figure supplement 2.**
Table of pairwise percent sequence identity of $3C^{pros}$ used in this study as determined from the alignment shown in Figure 4 – figure supplement 1.

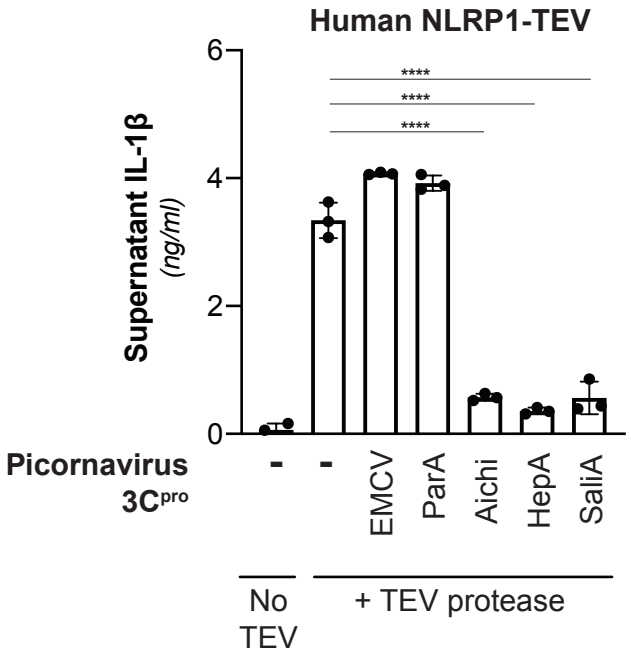


**Figure 4 – figure supplement 3.**
Inhibition of NLRP1 activation by non-enteroviral $3C^{pro}$. HEK293T cells were transfected with inflammasome components as in Figure 2E using 100ng TEV protease or pQCXIP empty vector, but with the additional inclusion of a non-enterovirus $3C^{pro}$ or empty vector (100ng). Release of bioactive IL-1β into the culture supernatant was measured using HEK-Blue IL-1β reporter cells as in Figure 2G. Data were analyzed using one-way ANOVA with Tukey's post-test comparing all conditions containing TEV protease. **** = $p < 0.0001$.

some degree (Figure 4B). For instance, while $3C^{pro}$ from EMCV and ParA did not cleave NLRP1, we observed distinct cleavage sites for $3C^{pro}$ from Aichi, HepA, SaliA and Rosa2 (Figure 4B), all of which have at least one cleavage site predicted to occur in the linker region (expected size between 40 kDa and 67 kDa). Confirming that these proteases cleave at a site that is distinct from that of enteroviruses, the G131P NLRP1 mutant is still cleaved by the non-enteroviral proteases (Figure 4B). To further confirm that cleavage (or lack thereof) of NLRP1 by expression of $3C^{pro}$ alone is reflective of $3C^{pro}$ during viral infection, we infected cells expressing WT or 131P NLRP1 with EMCV. Consistent with our co-transfection experiments, we see no cleavage of NLRP1 when we infect with EMCV, despite seeing robust cleavage when we infect with CVB3 (Figure 4C).

Surprisingly, when we interrogated NLRP1 inflammasome activation by $3C^{pros}$ from Aichi, HepA, SaliA and Rosa2, all of which robustly cleave NLRP1 at a site in the linker region, we found that only Rosa2 was able to activate the NLRP1 inflammasome (Figure 4D). While it is possible that NLRP1 cleavage by $3C^{pro}$ from these other viruses is too weak or in a region that may be inconsistent with activation, we also noted that there are obvious cleavage sites in NLRP1 that are outside of the linker region and closer to the FIIND autocleavage site. Cleavage at these sites in NLRP1, or cleavage of other host genes, may interfere with activation that may have otherwise been induced by $3C^{pro}$ cleavage in the linker region. Indeed, we find that co-expression of $3C^{pro}$ from Aichi, HepA, SaliA can attenuate NLRP1 activation by TEV protease (Figure 4 – figure supplement 3), consistent with the idea that these three proteases can actively block NLRP1 activation. Further investigation will be needed to determine the exact mechanism by which this occurs. Nevertheless, our data demonstrate that non-enteroviral $3C^{pros}$ can cleave NLRP1 at independent sites in the rapidly evolving linker region and can, in at least one case, activate the human NLRP1 inflammasome.

To further test the hypothesis that diversity of $3C^{pro}$ impacts host NLRP1 targeting, we compared the same $3C^{pro}$ panel against mouse NLRP1B (Figure 4E). Unlike $3C^{pro}$-mediated

cleavage of human NLRP1, in which all enteroviral $3C^{pro}$ cleave human NLRP1 at the same site, we found that even within enteroviruses there were differences in the ability to cleave NLRP1B and the cleavage sites. Outside of enteroviruses, we again found differences in sites cleaved within mouse NLRP1B. These data further support the model in which both host and viral evolution, even within closely related host and viral species, shape the outcome of the interaction between NLRP1 and $3C^{pro}$.

## DISCUSSION

Pathogens and their hosts are locked in a continual evolutionary conflict in which each side is attempting to exploit the others' weakness. One particularly successful strategy that pathogens have adopted is to exploit host processes that are highly constrained, leaving the host little room to evolutionarily adapt to overcome the pathogen. For instance, molecular mimicry of host proteins is commonly deployed by pathogens to antagonize host defenses, as it limits the evolutionary options for the host to counter-evolve (Elde & Malik, 2009). Beyond mimicry of entire proteins or protein domains, pathogens can also mimic so-called "short linear motifs" (SLIMs) through evolution of only a small number of amino acids to hijack highly conserved host processes such as post-translational modifications or binding by small protein domains (Chemes, de Prat-Gay, & Sanchez, 2015; Hagai, Azia, Babu, & Andino, 2014). Though these strategies are generally described as taking advantage of host evolutionary constraint, pathogens also have potential weak points of evolutionary constraint. In particular, proteases from positive-sense RNA viruses, such as picornaviruses, need to specifically cleave numerous sites within the viral polyprotein in order to reproduce. Thus, changing protease specificity requires concomitant changes to several independent cleavage sites, which is difficult to accomplish in a single evolutionary step. On top of that, protease cleavage motifs often only span a small number of amino acids (Schechter & Berger, 1967), potentially facilitating the independent evolution of these SLIMs in host proteins.

Here we show that the inflammasome protein, NLRP1, serves as a sensor for diverse proteases from the *Picornaviridae* family of human pathogens by mimicking the highly conserved protease cleavage sites found within the viral polyproteins. By exploiting a constrained feature of viral evolution and tying it to a pro-inflammatory immune response, such a system allows the immune system to recognize and respond to a wide range of viral proteases expressed in the host cytoplasm. NLRP1 represents one of the few known cases of

mammalian ETI (Cui et al., 2015; Fischer et al., 2020; Jones et al., 2016), where pathogen-mediated cleavage of NLRP1 promotes its activation. By holding the small C-terminal CARD-containing fragment in a non-covalent association with the larger N-terminal fragment, the majority of the protein can serve as a sensor for pathogen-encoded effectors (Mitchell et al., 2019; Taabazuing et al., 2020). This presents an opportunity to allow NLRP1 to evolve to be recognized by pathogenic effectors, ultimately leading to degradation of the N-terminal fragment. Indeed, mouse NLRP1B has been shown to be specifically cleaved by the protease-containing secreted effector from *B. anthracis* (LF) as well as a being ubiquitylated by an E3-ubiquitin ligase from *Shigella flexneri* (Sandstrom et al., 2019). While these two examples provide evidence that the mouse NLRP1B inflammasome operates by a "functional degradation" model, a direct pathogen-encoded activator of human NLRP1 has remained elusive. We now show, using an evolution-guided approach, that proteases from diverse picornaviruses, including human pathogens such as coxsackievirus B3 (CVB3), human rhinovirus A (HRVA), enterovirus D68 (EV68) and poliovirus 1 (PV1) and rosavirus A2 (Rosa2), specifically cleave several independently evolved sites in human NLRP1, leading to activation of the NLRP1 inflammasome and release of pro-inflammatory cytokines such as IL-1β. Thus, our work has identified viral proteases as the first known pathogen-encoded activators of human NLRP1.

We previously speculated that the unique domain architecture of NLRP1 would allow the N-terminal linker of human NLRP1 to freely evolve to be recognized by pathogenic effectors. Indeed, by harvesting publicly available enterovirus polyprotein sequences for known $3C^{pro}$ cleavage sites, we created a $3C^{pro}$ cleavage motif that was used to successfully predict the site of enterovirus $3C^{pro}$ cleavage at position 130-131 within the rapidly-evolving linker NLRP1. Additionally, our finding that numerous enteroviruses also cleave at the Q130-G131 site and activate pro-inflammatory cytokine release suggests that human NLRP1 serves as a general enteroviral protease sensor by encoding a polyprotein cleavage site mimic. Our phylogenetic assessment of the Q130-G131 $3C^{pro}$ cleavage site in NLRP1 suggests that NLRP1 sensing of

383 enteroviruses at this specific site is an innovation in the primate lineage, and is largely absent in
384 all other mammalian lineages with exception of a possible independent acquisition by members
385 within the *Caprinae* subfamily of mammals (*e.g.* goats, sheep) (Figure 3 – figure supplement 1).
386 Interestingly, even within the primate lineage and a small fraction of the human population,
387 some primate orthologs and human variants are cleavage-resistant and therefore do not
388 activate the inflammasome upon cytoplasmic expression of $3C^{pro}$. Such data may hint at three
389 different possible explanations for these changes. First, evolutionary drift in the absence of
390 pressure from pathogenic enteroviruses may account for loss of enterovirus $3C^{pro}$
391 responsiveness in these genes. Second, selection to sense another viral protease may shape
392 the same region of the linker. Finally, while the ETI model of NLRP1 suggests that enteroviral
393 cleavage of NLRP1 has evolved to activate a beneficial immune response in certain contexts,
394 the effects of NLRP1 overactivation may be detrimental in other contexts. In human skin
395 keratinocytes, where NLRP1 is regarded as the key inflammasome, all components of the
396 NLRP1 inflammasome are basally expressed and thus poised to elicit an inflammatory response
397 (Zhong et al., 2016). Here, germline mutations in NLRP1 that result in overactivation can cause
398 growth of warts in the upper airway in a condition known as recurrent respiratory papillomatosis
399 (JRRP) (Drutman et al., 2019) and an increase in skin cancer susceptibility and skin disorders
400 such as multiple self-healing palmoplantar carcinoma (MSPC), familial keratosis lichenoides
401 chronica (FKLC) and auto-inflammation with arthritis and dyskeratosis (AIADK) (Grandemange
402 et al., 2017; Herlin et al., 2019; Soler et al., 2013; Zhong et al., 2016; Zhong et al., 2018).
403 Beyond the skin, NLRP1 is also basally expressed in tissues such as the gut and brain
404 (D'Osualdo et al., 2015; Kaushal et al., 2015; Kummer et al., 2007), which are sites of
405 picornavirus replication where overactivation upon infection may result in immunopathology.
406 Further *in vivo* studies, facilitated by our discovery of specific picornavirus proteases that cleave
407 mouse NLRP1B, will help determine the role of NLRP1 in antiviral immunity and/or
408 immunopathology during viral infection.

Intriguingly, $3C^{pros}$ from nearly every genus of human-infecting picornavirus can cleave NLRP1 somewhere in the rapidly evolving linker region between the PYD and NLR domain, although only enteroviruses cleave at the specific site between position 130 and 131. These data suggest that this extended linker, which we previously found showed widespread signatures of positive selection (Chavarria-Smith et al., 2016), may be convergently evolving to mimic cleavage sites from a diverse range of viruses at multiple independent sites. Supporting that model, we observe a similar phenomenon in mouse NLRP1B, where multiple viruses cleave at different sites within NLRP1. These data highlight the important functional differences in cleavage specificity between even closely related $3C^{pro}$ that are not accounted for by predictive models. Further studies will be required to understand the precise relationships between sites within NLRP1 and individual protease specificity. Intriguingly, not all of these cleavage events lead to inflammasome activation in the same way that enteroviral cleavage does, and we find evidence for antagonism of NLRP1 activation by some $3C^{pros}$, suggesting that additional activities of $3C^{pro}$ may be the next step in the arms race, serving to prevent inflammasome activation even after the tripwire has been tripped.

Taken together, our work suggests that host mimicry of viral polyprotein cleavage motifs could be an important evolutionary strategy in the ongoing arms race between host and viruses. Indeed, one explanation for the somewhat surprising observation that the specificity of viral proteases changes at all within a viral family such as the picornaviruses is that there is evolutionary pressure from the host to evolve cleavage sites and protease specificity. Prior work has highlighted the roles that viral proteases can play in antagonizing host immune factors and driving host evolution to avoid being cleaved (Patel, Loo, Horner, Gale, & Malik, 2012; Stabell et al., 2018). In that case, the viral proteases would evolve to antagonize new factors while maintaining polyprotein cleavage. However, mimicry coupled with cleavage-activating immunity as seen with NLRP1 could be an even stronger pressure to shape the protease specificity. By turning the tables, these host processes may drive the type of functional diversification of viral

435 protease specificity as we observe in order to avoid cleaving NLRP1 and other similar ETI

436 factors. We expect that this work may lead to the discovery that such an evolutionary strategy

437 may be more broadly deployed at other sites of host-pathogen conflicts.

## MATERIALS AND METHODS

### Motif generation and search

To build the motif, 2658 nonredundant enteroviral polyprotein sequences were collected from the Viral Pathogen Resource (ViPR) and aligned with 20 well-annotated reference enteroviral polyprotein sequences from RefSeq (Supplementary file 1). P1 and P1' of the 8 annotated cleavage sites across the RefSeq sequences served as reference points for putative cleavage sites across the 2658 ViPR sequences, with the exception of enterovirus D polyproteins. The $3C^{pro}$ cleavage site for VP3-VP1 within polyproteins from the clade of enterovirus D have been described to be undetectable and have thus been removed (Tan et al., 2013). Four amino acyl residues upstream (P4-P1) and downstream (P1'-P4') of each cleavage site were extracted from every MAFFT-aligned polyprotein sequence, resulting in 2678 sets of cleavage sites (RefSeq sites included). Each set of cleavage sites representative of each polyprotein was then concatenated. 1884 duplicates were removed from the 2678 concatenated cleavage sites. The remaining 796 nonredundant, concatenated cleavage sites were then split into individual 8-mer cleavage sites and the 6333 8-mers were aligned using MAFFT to generate Geneious-defined sequence logo information at each aligned position. Pseudo-counts to the position-specific scoring matrix were adjusted by total information content within each position relative to the two most information-dense position P1 and P1' (pseudocount = 0) and the least information-dense position P3 (pseudocount = 1). The 0.002 p-value threshold for FIMO motif searching against human NLRP1 was determined to optimize the capture of 95% of initial input cleavage sites within the set of 2678 whole enteroviral polyproteins and a majority sites within a previously described dataset of enteroviral $3C^{pro}$ targets (Laitinen et al., 2016).

### NetSurfP

Prediction of the coil probability across human NLRP1 (NCBI accession NP_127497.1) was conducted using the protein FASTA as the input for the NetSurfP web server (http://www.cbs.dtu.dk/services/NetSurfP/).

**Sequence alignments, phylogenetic trees and NLRP1 phylogenomics**

Complete polyprotein sequences from 796 picornaviruses with non-redundant $3C^{pro}$ cleavage sites (see "Motif generation and search" section above) were downloaded from ViPR. Sequences were aligned using MAFFT (Katoh & Standley, 2013) and a neighbor-joining phylogenetic tree was generated using Geneious software (Kearse et al., 2012). An alignment and phylogenetic tree of all of the $3C^{pro}$ sequences used in this study was generated similarly. To identify mammalian NLRP1 homologs, and species that lack NLRP1, the human NLRP1 protein sequence was used to query the RefSeq protein sequence database, a curated collection of the most well-assembled genomes, using BLASTp (Altschul et al., 1997). Sequences were downloaded and aligned using MAFFT implemented in Geneious software. Consensus sequence logos shown were generated using Geneious software. We determined that NLRP1 was "absent" from a clade of species using the following criteria: (1) when searching with human NLRP1, we found an obvious homolog of another NLRP protein (generally NLRP3, NLRP12 or NLRP14) but no complete or partial homolog of NLRP1 and (2) this absence was apparent in every member of the clade of species (>2 species) in the RefSeq database.

**Plasmids and constructs**

For NLRP1 cleavage assays, the coding sequences of human NLRP1 WT, human NLRP1 mutants (G131P, G131R, Q130R, G127E), human NLRP1 TEV or mouse NLRP1B were cloned into the pcDNA5/FRT/TO backbone (Invitrogen, Carlsbad, CA) with an N-terminal 3xFlag and mCherry tag. For NLRP1 activation, the coding sequences of human NLRP1 WT (NCBI

accession NP_127497.1), human NLRP1 mutants (G131P, G131R, Q130R, G127E), human NLRP1 TEV or mouse NLRP1B (mouse strain 129 allele, NCBI accession AAZ40510.1) were cloned into the pQCXIP vector backbone (Takara Bio, Mountain View, CA) with a C-terminal Myc tag. Vectors containing the coding sequences of human NLRP1 TEV (NLRP1-TEV2), ASC, CASP1, IL-1β-V5, and TEV protease (Chavarria-Smith et al., 2016) were generous gifts from Dr. Russell Vance, UC Berkeley. Single point mutations were made using overlapping stitch PCR. A list of primers used to generate the wild-type and mutant NLRP1 constructs are described in Supplementary file 5.

CVB3 $3C^{pro}$ and EMCV $3C^{pro}$ were cloned from CVB3-Nancy and EMCV-Mengo plasmids (generous gifts from Dr. Julie Pfeiffer, UT Southwestern). Remaining $3C^{pro}$ sequences were ordered as gBlocks (Integrated DNA Technologies, San Diego, CA). Each $3C^{pro}$ was cloned with an N-terminal HA tag into the QCXIP vector backbone. Catalytic mutations were made using overlapping stitch PCR. A list of primers and gBlocks used to generate the protease constructs are described in Supplementary file 5.

Following cloning, all plasmid stocks were sequenced across the entire inserted region to verify that no mutations were introduced during the cloning process.

**Cell culture and transient transfection**

All cells are routinely tested for mycoplasma by PCR kit (ATCC, Manassas, VA) and grown in complete media containing DMEM (Gibco, Carlsbad, CA), 10% FBS (Peak Serum, Wellington, CO), and appropriate antibiotics (Gibco, Carlsbad, CA). For transient transfections, HEK293T cells were seeded the day prior to transfection in a 24-well plate (Genesee, El Cajon, CA) with 500µl complete media. Cells were transiently transfected with 500ng of total DNA and 1.5µl of Transit X2 (Mirus Bio, Madison, WI) following the manufacturer's protocol. HEK-Blue IL-1β

reporter cells (Invivogen, San Diego, CA) were grown and assayed in 96-well plates (Genesee, El Cajon, CA).

**NLRP1 cleavage assays**

100ng of human NLRP1 WT, human NLRP1 mutants (G131P, G131R, Q130R, G127E), human NLRP1 TEV or mouse NLRP1B tagged with mCherry, 3xFLAG and Myc was co-transfected with 250ng of HA-tagged protease-producing constructs. Twenty-four hours post-transfection, the cells were harvested, lysed in 1x NuPAGE LDS sample buffer (Invitrogen, Carlsbad, CA) containing 5% β-mercaptoethanol (Fisher Scientific, Pittsburg, PA) and immunoblotted with antibodies described below.

**Viral stocks and viral infections**

CVB3 and EMCV viral stocks were generated by co-transfection of CVB3-Nancy or EMCV-Mengo infectious clone plasmids with a plasmid expressing T7 RNA polymerase (generous gifts from Dr. Julie Pfeiffer, UT Southwestern) as previously described (McCune, Lanahan, tenOever, & Pfeiffer, 2020). Supernatant was harvested, quantified by TCID50 on HEK293T cells, and frozen in aliquots at -80°C.

For viral infections, cells in 24-well plates were transfected with the indicated NLRP1 plasmid and infected with 250,000 PFU (MOI = ~1) of the indicated virus for the indicated times. All samples were collected 32 hours after transfection. The cells were harvested, lysed in 1x NuPAGE LDS sample buffer (Invitrogen, Carlsbad, CA) containing 5% β-mercaptoethanol (Fisher Scientific, Pittsburg, PA) and immunoblotted with antibodies described below.

**NLRP1 activity assays**

5ng of ASC, 100ng of CASP1, 50ng of IL-1β-V5, and 100ng of various protease-producing constructs were co-transfected with 4ng of either pQCXIP empty vector, wild-type or mutant pQCXIP-NLRP1-Myc constructs. Twenty-four hours post-transfection, the cells were harvested and lysed in 1x NuPAGE LDS sample buffer containing 5% β-mercaptoethanol and immunoblotted with antibodies described below or culture media was harvested for quantification of IL-1β levels by HEK-Blue assays (see below).

**HEK-Blue IL-1β assay**

To quantify the levels of bioactive IL-1β released from cells, we employed HEK-Blue IL-1β reporter cells. In these cells, binding to IL-1β to the surface receptor IL-1R1 results in the downstream activation of NF-kB and subsequent production of secreted embryonic alkaline phosphatase (SEAP) in a dose-dependent manner. SEAP levels are detected using a colorimetric substrate assay, QUANTI-Blue (Invivogen, San Diego, CA) by measuring an increase in absorbance at $OD_{655}$.

Twenty-four hours after transfection of HEK293T cells with inflammasome components and $3C^{pro}$ (see NLRP1 activity assays above), 10µl of the culture supernatants was added to HEK-Blue IL-1β reporter cells plated in 96-well format in a total volume of 200µl per well. On the same plate, serial dilutions of recombinant human IL-1β (Invivogen, San Diego, CA) were added in order to generate a standard curve for each assay. Twenty-four hours later, SEAP levels were assayed by taking 20µl of the supernatant from HEK-Blue IL-1β reporter cells and adding to 180µl of QUANTI-Blue colorimetric substrate following the manufacturer's protocol. After incubation at 37°C for 30-60 minutes, absorbance at $OD_{655}$ was measured on a BioTek Cytation 5 plate reader (BioTek Instruments, Winooski, VT) and absolute levels of IL-1β were calculated

564 relative to the standard curve. All assays, beginning with independent transfections, were
565 performed in triplicate.
566
567 **Immunoblotting and antibodies**
568 Harvested cell pellets were washed with 1X PBS, and lysed with 1x NuPAGE LDS sample
569 buffer containing 5% β-mercaptoethanol at 98C for 10 minutes. The lysed samples were spun
570 down at 15000 RPM for two minutes, followed by loading into a 4-12% Bis-Tris SDS-PAGE gel
571 (Life Technologies, San Diego, CA) with 1X MOPS buffer (Life Technologies, San Diego, CA)
572 and wet transfer onto a nitrocellulose membrane (Life Technologies, San Diego, CA).
573 Membranes were blocked with PBS-T containing 5% bovine serum albumin (BSA) (Spectrum,
574 New Brunswick, NJ), followed by incubation with primary antibodies for V5 (IL-1β), FLAG
575 (mCherry-fused NLRP1 for protease assays), Myc (NLRP1-Myc for activation assays), HA (viral
576 protease), or GAPDH. Membranes were rinsed three times in PBS-T then incubated with the
577 appropriate HRP-conjugated secondary antibodies. Membranes were rinsed again three times
578 in PBS-T and developed with SuperSignal West Pico PLUS Chemiluminescent Substrate
579 (Thermo Fisher Scientific, Carlsbad, CA). The specifications, source and clone info for
580 antibodies are described in Supplementary file 6.

## ACKNOWLEDGEMENTS

We thank members of the Daugherty laboratory and members of the San Diego Program in Immunology for helpful suggestions and Andrew Sandstrom and members of the Daugherty laboratory for critical reading of the manuscript. This work was supported by the National Institutes of Health (R35 GM133633), Pew Biomedical Scholars Program and Hellman Fellows Program to M.D.D, T32 GM007240 to B.V.T., C.B. and A.P.R, a National Science Foundation graduate research fellowship (2019284620) to C.B., and a Jane Coffin Childs Memorial Fund for Medical Research postdoctoral fellowship to P.S.M.

590 **REFERENCES**

591

592 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J.
593 (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search
594 programs. *Nucleic Acids Res, 25*(17), 3389-3402. doi:10.1093/nar/25.17.3389

595 Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., . . . Noble, W. S.
596 (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res,*
597 *37*(Web Server issue), W202-208. doi:10.1093/nar/gkp335

598 Blom, N., Hansen, J., Blaas, D., & Brunak, S. (1996). Cleavage site analysis in picornaviral
599 polyproteins: discovering cellular targets by neural networks. *Protein Sci, 5*(11), 2203-
600 2216. doi:10.1002/pro.5560051107

601 Boyden, E. D., & Dietrich, W. F. (2006). Nalp1b controls mouse macrophage susceptibility to
602 anthrax lethal toxin. *Nat Genet, 38*(2), 240-244. doi:10.1038/ng1724

603 Broz, P., & Dixit, V. M. (2016). Inflammasomes: mechanism of assembly, regulation and
604 signalling. *Nat Rev Immunol, 16*(7), 407-420. doi:10.1038/nri.2016.58

605 Cagliani, R., Forni, D., Tresoldi, C., Pozzoli, U., Filippi, G., Rainone, V., . . . Sironi, M. (2014).
606 RIG-I-like receptors evolved adaptively in mammals, with parallel evolution at LGP2 and
607 RIG-I. *J Mol Biol, 426*(6), 1351-1365. doi:10.1016/j.jmb.2013.10.040

608 Chavarria-Smith, J., Mitchell, P. S., Ho, A. M., Daugherty, M. D., & Vance, R. E. (2016).
609 Functional and Evolutionary Analyses Identify Proteolysis as a General Mechanism for
610 NLRP1 Inflammasome Activation. *PLoS Pathog, 12*(12), e1006052.
611 doi:10.1371/journal.ppat.1006052

612 Chavarria-Smith, J., & Vance, R. E. (2013). Direct proteolytic cleavage of NLRP1B is necessary
613 and sufficient for inflammasome activation by anthrax lethal factor. *PLoS Pathog, 9*(6),
614 e1003452. doi:10.1371/journal.ppat.1003452

615 Chemes, L. B., de Prat-Gay, G., & Sanchez, I. E. (2015). Convergent evolution and mimicry of
616 protein linear motifs in host-pathogen interactions. *Curr Opin Struct Biol, 32*, 91-101.
617 doi:10.1016/j.sbi.2015.03.004
618 Chui, A. J., Okondo, M. C., Rao, S. D., Gai, K., Griswold, A. R., Johnson, D. C., . . . Bachovchin,
619 D. A. (2019). N-terminal degradation activates the NLRP1B inflammasome. *Science,*
620 *364*(6435), 82-85. doi:10.1126/science.aau1208
621 Croft, S. N., Walker, E. J., & Ghildyal, R. (2018). Human Rhinovirus 3C protease cleaves
622 RIPK1, concurrent with caspase 8 activation. *Sci Rep, 8*(1), 1569. doi:10.1038/s41598-
623 018-19839-4
624 Cui, H., Tsuda, K., & Parker, J. E. (2015). Effector-triggered immunity: from pathogen
625 perception to robust defense. *Annu Rev Plant Biol, 66*, 487-511. doi:10.1146/annurev-
626 arplant-050213-040012
627 D'Osualdo, A., Anania, V. G., Yu, K., Lill, J. R., Kaufman, R. J., Matsuzawa, S., & Reed, J. C.
628 (2015). Transcription Factor ATF4 Induces NLRP1 Inflammasome Expression during
629 Endoplasmic Reticulum Stress. *PLoS One, 10*(6), e0130635.
630 doi:10.1371/journal.pone.0130635
631 D'Osualdo, A., Weichenberger, C. X., Wagner, R. N., Godzik, A., Wooley, J., & Reed, J. C.
632 (2011). CARD8 and NLRP1 undergo autoproteolytic processing through a ZU5-like
633 domain. *PLoS One, 6*(11), e27396. doi:10.1371/journal.pone.0027396
634 Daugherty, M. D., & Malik, H. S. (2012). Rules of engagement: molecular insights from host-
635 virus arms races. *Annu Rev Genet, 46*, 677-700. doi:10.1146/annurev-genet-110711-
636 155522
637 Drutman, S. B., Haerynck, F., Zhong, F. L., Hum, D., Hernandez, N. J., Belkaya, S., . . .
638 Casanova, J. L. (2019). Homozygous NLRP1 gain-of-function mutation in siblings with a
639 syndromic form of recurrent respiratory papillomatosis. *Proc Natl Acad Sci U S A,*
640 *116*(38), 19055-19063. doi:10.1073/pnas.1906184116

641 Elde, N. C., & Malik, H. S. (2009). The evolutionary conundrum of pathogen mimicry. *Nat Rev*
642 *Microbiol, 7*(11), 787-797. doi:10.1038/nrmicro2222
643 Evavold, C. L., & Kagan, J. C. (2019). Inflammasomes: Threat-Assessment Organelles of the
644 Innate Immune System. *Immunity, 51*(4), 609-624. doi:10.1016/j.immuni.2019.08.005
645 Fan, X., Li, X., Zhou, Y., Mei, M., Liu, P., Zhao, J., . . . Yi, L. (2020). Quantitative Analysis of the
646 Substrate Specificity of Human Rhinovirus 3C Protease and Exploration of Its Substrate
647 Recognition Mechanisms. *ACS Chem Biol, 15*(1), 63-73.
648 doi:10.1021/acschembio.9b00539
649 Finger, J. N., Lich, J. D., Dare, L. C., Cook, M. N., Brown, K. K., Duraiswami, C., . . . Gough, P.
650 J. (2012). Autolytic proteolysis within the function to find domain (FIIND) is required for
651 NLRP1 inflammasome activity. *J Biol Chem, 287*(30), 25030-25037.
652 doi:10.1074/jbc.M112.378323
653 Fischer, N. L., Naseer, N., Shin, S., & Brodsky, I. E. (2020). Publisher Correction: Effector-
654 triggered immunity and pathogen sensing in metazoans. *Nat Microbiol, 5*(3), 528.
655 doi:10.1038/s41564-020-0682-4
656 Frew, B. C., Joag, V. R., & Mogridge, J. (2012). Proteolytic processing of Nlrp1b is required for
657 inflammasome activity. *PLoS Pathog, 8*(4), e1002659. doi:10.1371/journal.ppat.1002659
658 Grandemange, S., Sanchez, E., Louis-Plence, P., Tran Mau-Them, F., Bessis, D., Coubes, C., .
659 . . Genevieve, D. (2017). A new autoinflammatory and autoimmune syndrome
660 associated with NLRP1 mutations: NAIAD (NLRP1-associated autoinflammation with
661 arthritis and dyskeratosis). *Ann Rheum Dis, 76*(7), 1191-1198.
662 doi:10.1136/annrheumdis-2016-210021
663 Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: scanning for occurrences of a given
664 motif. *Bioinformatics, 27*(7), 1017-1018. doi:10.1093/bioinformatics/btr064
665 Greaney, A. J., Portley, M. K., O'Mard, D., Crown, D., Maier, N. K., Mendenhall, M. A., . . .
666 Moayeri, M. (2020). Frontline Science: Anthrax lethal toxin-induced, NLRP1-mediated IL-

1beta release is a neutrophil and PAD4-dependent event. *J Leukoc Biol*. doi:10.1002/JLB.4HI0320-028R

Hagai, T., Azia, A., Babu, M. M., & Andino, R. (2014). Use of host-like peptide motifs in viral proteins is a prevalent strategy in host-virus interactions. *Cell Rep, 7*(5), 1729-1739. doi:10.1016/j.celrep.2014.04.052

Hancks, D. C., Hartley, M. K., Hagan, C., Clark, N. L., & Elde, N. C. (2015). Overlapping Patterns of Rapid Evolution in the Nucleic Acid Sensors cGAS and OAS1 Suggest a Common Mechanism of Pathogen Antagonism and Escape. *PLoS Genet, 11*(5), e1005203. doi:10.1371/journal.pgen.1005203

Herlin, T., Jorgensen, S. E., Host, C., Mitchell, P. S., Christensen, M. H., Laustsen, M., . . . Mogensen, T. H. (2019). Autoinflammatory disease with corneal and mucosal dyskeratosis caused by a novel NLRP1 variant. *Rheumatology (Oxford)*. doi:10.1093/rheumatology/kez612

Huang, L., Liu, Q., Zhang, L., Zhang, Q., Hu, L., Li, C., . . . Weng, C. (2015). Encephalomyocarditis Virus 3C Protease Relieves TRAF Family Member-associated NF-kappaB Activator (TANK) Inhibitory Effect on TRAF6-mediated NF-kappaB Signaling through Cleavage of TANK. *J Biol Chem, 290*(46), 27618-27632. doi:10.1074/jbc.M115.660761

Jagdeo, J. M., Dufour, A., Klein, T., Solis, N., Kleifeld, O., Kizhakkedathu, J., . . . Jan, E. (2018). N-Terminomics TAILS Identifies Host Cell Substrates of Poliovirus and Coxsackievirus B3 3C Proteinases That Modulate Virus Infection. *J Virol, 92*(8). doi:10.1128/JVI.02211-17

Janeway, C. A., Jr. (1989). Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb Symp Quant Biol, 54 Pt 1*, 1-13. doi:10.1101/sqb.1989.054.01.003

692 Jones, J. D., Vance, R. E., & Dangl, J. L. (2016). Intracellular innate immune surveillance
693 devices in plants and animals. *Science, 354*(6316). doi:10.1126/science.aaf6395
694 Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alfoldi, J., Wang, Q., . . .
695 MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in
696 141,456 humans. *Nature, 581*(7809), 434-443. doi:10.1038/s41586-020-2308-7
697 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7:
698 improvements in performance and usability. *Mol Biol Evol, 30*(4), 772-780.
699 doi:10.1093/molbev/mst010
700 Kaushal, V., Dye, R., Pakavathkumar, P., Foveau, B., Flores, J., Hyman, B., . . . LeBlanc, A. C.
701 (2015). Neuronal NLRP1 inflammasome activation of Caspase-1 coordinately regulates
702 inflammatory interleukin-1-beta production and axonal degeneration-associated
703 Caspase-6 activation. *Cell Death Differ, 22*(10), 1676-1686. doi:10.1038/cdd.2015.16
704 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., . . . Drummond, A.
705 (2012). Geneious Basic: an integrated and extendable desktop software platform for the
706 organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649.
707 doi:10.1093/bioinformatics/bts199
708 Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge Cambridgeshire ; New
709 York: Cambridge University Press.
710 Klausen, M. S., Jespersen, M. C., Nielsen, H., Jensen, K. K., Jurtz, V. I., Sonderby, C. K., . . .
711 Marcatili, P. (2019). NetSurfP-2.0: Improved prediction of protein structural features by
712 integrated deep learning. *Proteins, 87*(6), 520-527. doi:10.1002/prot.25674
713 Kummer, J. A., Broekhuizen, R., Everett, H., Agostini, L., Kuijk, L., Martinon, F., . . . Tschopp, J.
714 (2007). Inflammasome components NALP 1 and 3 show distinct but separate expression
715 profiles in human tissues suggesting a site-specific role in the inflammatory response. *J*
716 *Histochem Cytochem, 55*(5), 443-452. doi:10.1369/jhc.6A7101.2006

717 Laitinen, O. H., Svedin, E., Kapell, S., Nurminen, A., Hytonen, V. P., & Flodstrom-Tullberg, M.
718 (2016). Enteroviral proteases: structure, host interactions and pathogenicity. *Rev Med*
719 *Virol, 26*(4), 251-267. doi:10.1002/rmv.1883
720 Lei, X., Zhang, Z., Xiao, X., Qi, J., He, B., & Wang, J. (2017). Enterovirus 71 Inhibits Pyroptosis
721 through Cleavage of Gasdermin D. *J Virol, 91*(18). doi:10.1128/JVI.01069-17
722 Levinsohn, J. L., Newman, Z. L., Hellmich, K. A., Fattah, R., Getz, M. A., Liu, S., . . . Moayeri, M.
723 (2012). Anthrax lethal factor cleavage of Nlrp1 is required for activation of the
724 inflammasome. *PLoS Pathog, 8*(3), e1002638. doi:10.1371/journal.ppat.1002638
725 Martinon, F., Burns, K., & Tschopp, J. (2002). The inflammasome: a molecular platform
726 triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell,*
727 *10*(2), 417-426. doi:10.1016/s1097-2765(02)00599-3
728 McCune, B. T., Lanahan, M. R., tenOever, B. R., & Pfeiffer, J. K. (2020). Rapid Dissemination
729 and Monopolization of Viral Populations in Mice Revealed Using a Panel of Barcoded
730 Viruses. *J Virol, 94*(2). doi:10.1128/JVI.01590-19
731 Meyerson, N. R., & Sawyer, S. L. (2011). Two-stepping through time: mammals and viruses.
732 *Trends Microbiol, 19*(6), 286-294. doi:10.1016/j.tim.2011.03.006
733 Mitchell, P. S., Sandstrom, A., & Vance, R. E. (2019). The NLRP1 inflammasome: new
734 mechanistic insights and unresolved mysteries. *Curr Opin Immunol, 60*, 37-45.
735 doi:10.1016/j.coi.2019.04.015
736 Moayeri, M., Crown, D., Newman, Z. L., Okugawa, S., Eckhaus, M., Cataisson, C., . . . Leppla,
737 S. H. (2010). Inflammasome sensor Nlrp1b-dependent resistance to anthrax is mediated
738 by caspase-1, IL-1 signaling and neutrophil recruitment. *PLoS Pathog, 6*(12), e1001222.
739 doi:10.1371/journal.ppat.1001222
740 Moayeri, M., Sastalla, I., & Leppla, S. H. (2012). Anthrax and the inflammasome. *Microbes*
741 *Infect, 14*(5), 392-400. doi:10.1016/j.micinf.2011.12.005

742 Mukherjee, A., Morosky, S. A., Delorme-Axford, E., Dybdahl-Sissoko, N., Oberste, M. S., Wang,
743 T., & Coyne, C. B. (2011). The coxsackievirus B 3C protease cleaves MAVS and TRIF
744 to attenuate host type I interferon and apoptotic signaling. *PLoS Pathog, 7*(3), e1001311.
745 doi:10.1371/journal.ppat.1001311
746 O'Donoghue, A. J., Eroy-Reveles, A. A., Knudsen, G. M., Ingram, J., Zhou, M., Statnekov, J. B.,
747 . . . Craik, C. S. (2012). Global identification of peptidase specificity by multiplex
748 substrate profiling. *Nat Methods, 9*(11), 1095-1100. doi:10.1038/nmeth.2182
749 Patel, M. R., Loo, Y. M., Horner, S. M., Gale, M., Jr., & Malik, H. S. (2012). Convergent
750 evolution of escape from hepaciviral antagonism in primates. *PLoS Biol, 10*(3),
751 e1001282. doi:10.1371/journal.pbio.1001282
752 Pickett, B. E., Greer, D. S., Zhang, Y., Stewart, L., Zhou, L., Sun, G., . . . Scheuermann, R. H.
753 (2012). Virus pathogen database and analysis resource (ViPR): a comprehensive
754 bioinformatics database and analysis resource for the coronavirus research community.
755 *Viruses, 4*(11), 3209-3226. doi:10.3390/v4113209
756 Qian, S., Fan, W., Liu, T., Wu, M., Zhang, H., Cui, X., . . . Qian, P. (2017). Seneca Valley Virus
757 Suppresses Host Type I Interferon Production by Targeting Adaptor Proteins MAVS,
758 TRIF, and TANK for Cleavage. *J Virol, 91*(16). doi:10.1128/JVI.00823-17
759 Rathinam, V. A., & Fitzgerald, K. A. (2016). Inflammasome Complexes: Emerging Mechanisms
760 and Effector Functions. *Cell, 165*(4), 792-800. doi:10.1016/j.cell.2016.03.046
761 Rathinam, V. A., Vanaja, S. K., & Fitzgerald, K. A. (2012). Regulation of inflammasome
762 signaling. *Nat Immunol, 13*(4), 333-342. doi:10.1038/ni.2237
763 Sandstrom, A., Mitchell, P. S., Goers, L., Mu, E. W., Lesser, C. F., & Vance, R. E. (2019).
764 Functional degradation: A mechanism of NLRP1 inflammasome activation by diverse
765 pathogen enzymes. *Science, 364*(6435). doi:10.1126/science.aau1330
766 Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain.
767 *Biochem Biophys Res Commun, 27*(2), 157-162. doi:10.1016/s0006-291x(67)80055-x

768 Sironi, M., Cagliani, R., Forni, D., & Clerici, M. (2015). Evolutionary insights into host-pathogen
769 interactions from mammalian sequence data. *Nat Rev Genet, 16*(4), 224-236.
770 doi:10.1038/nrg3905
771 Soler, V. J., Tran-Viet, K. N., Galiacy, S. D., Limviphuvadh, V., Klemm, T. P., St Germain, E., . .
772 . Young, T. L. (2013). Whole exome sequencing identifies a mutation for a novel form of
773 corneal intraepithelial dyskeratosis. *J Med Genet, 50*(4), 246-254.
774 doi:10.1136/jmedgenet-2012-101325
775 Solomon, T., Lewthwaite, P., Perera, D., Cardosa, M. J., McMinn, P., & Ooi, M. H. (2010).
776 Virology, epidemiology, pathogenesis, and control of enterovirus 71. *Lancet Infect Dis,*
777 *10*(11), 778-790. doi:10.1016/S1473-3099(10)70194-8
778 Stabell, A. C., Meyerson, N. R., Gullberg, R. C., Gilchrist, A. R., Webb, K. J., Old, W. M., . . .
779 Sawyer, S. L. (2018). Dengue viruses cleave STING in humans but not in nonhuman
780 primates, their presumed natural reservoir. *Elife, 7*. doi:10.7554/eLife.31919
781 Sun, D., Chen, S., Cheng, A., & Wang, M. (2016). Roles of the Picornaviral 3C Proteinase in the
782 Viral Life Cycle and Host Cells. *Viruses, 8*(3), 82. doi:10.3390/v8030082
783 Taabazuing, C. Y., Griswold, A. R., & Bachovchin, D. A. (2020). The NLRP1 and CARD8
784 inflammasomes. *Immunol Rev*. doi:10.1111/imr.12884
785 Tan, J., George, S., Kusov, Y., Perbandt, M., Anemuller, S., Mesters, J. R., . . . Hilgenfeld, R.
786 (2013). 3C protease of enterovirus 68: structure-based design of Michael acceptor
787 inhibitors and their broad-spectrum antiviral effects against picornaviruses. *J Virol, 87*(8),
788 4339-4351. doi:10.1128/JVI.01123-12
789 Tenthorey, J. L., Kofoed, E. M., Daugherty, M. D., Malik, H. S., & Vance, R. E. (2014). Molecular
790 basis for specific recognition of bacterial ligands by NAIP/NLRC4 inflammasomes. *Mol*
791 *Cell, 54*(1), 17-29. doi:10.1016/j.molcel.2014.02.018
792 Terra, J. K., Cote, C. K., France, B., Jenkins, A. L., Bozue, J. A., Welkos, S. L., . . . Bradley, K.
793 A. (2010). Cutting edge: resistance to Bacillus anthracis infection mediated by a lethal

toxin sensitive allele of Nalp1b/Nlrp1b. *J Immunol, 184*(1), 17-20. doi:10.4049/jimmunol.0903114

Tian, X., Pascal, G., & Monget, P. (2009). Evolution and functional divergence of NLRP genes in mammalian reproductive systems. *BMC Evol Biol, 9*, 202. doi:10.1186/1471-2148-9-202

Ting, J. P., Lovering, R. C., Alnemri, E. S., Bertin, J., Boss, J. M., Davis, B. K., . . . Ward, P. A. (2008). The NLR gene family: a standard nomenclature. *Immunity, 28*(3), 285-287. doi:10.1016/j.immuni.2008.02.005

Vance, R. E., Isberg, R. R., & Portnoy, D. A. (2009). Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. *Cell Host Microbe, 6*(1), 10-21. doi:10.1016/j.chom.2009.06.007

Wang, C., Fung, G., Deng, H., Jagdeo, J., Mohamud, Y., Xue, Y. C., . . . Luo, H. (2019). NLRP3 deficiency exacerbates enterovirus infection in mice. *FASEB J, 33*(1), 942-952. doi:10.1096/fj.201800301RRR

Wang, D., Fang, L., Li, K., Zhong, H., Fan, J., Ouyang, C., . . . Xiao, S. (2012). Foot-and-mouth disease virus 3C protease cleaves NEMO to impair innate immune signaling. *J Virol, 86*(17), 9311-9322. doi:10.1128/JVI.00722-12

Wang, D., Fang, L., Wei, D., Zhang, H., Luo, R., Chen, H., . . . Xiao, S. (2014). Hepatitis A virus 3C protease cleaves NEMO to impair induction of beta interferon. *J Virol, 88*(17), 10252-10258. doi:10.1128/JVI.00869-14

Wang, H., Lei, X., Xiao, X., Yang, C., Lu, W., Huang, Z., . . . Wang, J. (2015). Reciprocal Regulation between Enterovirus 71 and the NLRP3 Inflammasome. *Cell Rep, 12*(1), 42-48. doi:10.1016/j.celrep.2015.05.047

Wen, W., Yin, M., Zhang, H., Liu, T., Chen, H., Qian, P., . . . Li, X. (2019). Seneca Valley virus 2C and 3C inhibit type I interferon production by inducing the degradation of RIG-I. *Virology, 535*, 122-129. doi:10.1016/j.virol.2019.06.017

820 Wickliffe, K. E., Leppla, S. H., & Moayeri, M. (2008). Killing of macrophages by anthrax lethal
821 toxin: involvement of the N-end rule pathway. *Cell Microbiol, 10*(6), 1352-1362.
822 doi:10.1111/j.1462-5822.2008.01131.x
823 Xiang, Z., Li, L., Lei, X., Zhou, H., Zhou, Z., He, B., & Wang, J. (2014). Enterovirus 68 3C
824 protease cleaves TRIF to attenuate antiviral responses mediated by Toll-like receptor 3.
825 *J Virol, 88*(12), 6650-6659. doi:10.1128/JVI.03138-13
826 Xiang, Z., Liu, L., Lei, X., Zhou, Z., He, B., & Wang, J. (2016). 3C Protease of Enterovirus D68
827 Inhibits Cellular Defense Mediated by Interferon Regulatory Factor 7. *J Virol, 90*(3),
828 1613-1621. doi:10.1128/JVI.02395-15
829 Xu, H., Shi, J., Gao, H., Liu, Y., Yang, Z., Shao, F., & Dong, N. (2019). The N-end rule ubiquitin
830 ligase UBR2 mediates NLRP1B inflammasome activation by anthrax lethal toxin. *EMBO*
831 *J, 38*(13), e101996. doi:10.15252/embj.2019101996
832 Zaragoza, C., Saura, M., Padalko, E. Y., Lopez-Rivera, E., Lizarbe, T. R., Lamas, S., &
833 Lowenstein, C. J. (2006). Viral protease cleavage of inhibitor of kappaBalpha triggers
834 host cell apoptosis. *Proc Natl Acad Sci U S A, 103*(50), 19051-19056.
835 doi:10.1073/pnas.0606019103
836 Zell, R. (2018). Picornaviridae-the ever-growing virus family. *Arch Virol, 163*(2), 299-317.
837 doi:10.1007/s00705-017-3614-8
838 Zhong, F. L., Mamai, O., Sborgi, L., Boussofara, L., Hopkins, R., Robinson, K., . . . Reversade,
839 B. (2016). Germline NLRP1 Mutations Cause Skin Inflammatory and Cancer
840 Susceptibility Syndromes via Inflammasome Activation. *Cell, 167*(1), 187-202 e117.
841 doi:10.1016/j.cell.2016.09.001
842 Zhong, F. L., Robinson, K., Teo, D. E. T., Tan, K. Y., Lim, C., Harapas, C. R., . . . Reversade, B.
843 (2018). Human DPP9 represses NLRP1 inflammasome and protects against
844 autoinflammatory diseases via both peptidase activity and FIIND domain binding. *J Biol*
845 *Chem, 293*(49), 18864-18878. doi:10.1074/jbc.RA118.004350

846

847

## FIGURE LEGENDS

**Figure 1. Conserved polyprotein cleavage sites across enteroviruses inform substrate specificity of the enteroviral 3C$^{pro}$.** (A) Schematic of 3C$^{pro}$ cleavage sites (red arrows) within the polyprotein of coxsackievirus B3 Nancy (CVB3), a model enterovirus. Shown are the eight amino acids flanking each cleavage site within the polyprotein. (B) Phylogenetic tree of 796 enteroviral polyprotein coding sequences depicting the major clades of enteroviruses sampled in this study with representative viruses from each clade in parentheses (Supplementary file 2). (C) Eight amino acid polyprotein cleavage motif for enteroviruses (labeled as positions P4 to P4') generated from the 796 enteroviral polyprotein sequences in (B) using the MEME Suite (Supplementary file 2). (D) Training set data used to determine the motif search threshold for FIMO (Supplementary files 1, 3 and 4). The X-axis represents a $\log_{10}$ of the p-value reported by FIMO as an indicator for the strength of the cleavage motif hit (cleavage score). (Left) The Y-axis depicts the number of uncalled true positives, or motif hits that overlap with the initial set of 8mer polyprotein cleavage sites used to generate the motif, in the training set of enteroviral polyprotein sequences (black). (Right) The Y-axis depicts the number of called false positive sites, or any motif hits found in the polyprotein that are not known to be cleaved by 3C$^{pro}$, in the training set of enteroviral polyprotein sequences (gray). (Above) Each line depicts a single, experimentally validated case of enteroviral 3C$^{pro}$ cleavage site within a human protein as reported in Laitinen et al, 2016 and is ordered along the x-axis by its resulting cleavage score. A vertical dotted line is used to represent the decided threshold that captures 95% of true positive hits and 16 out of 27 reported human hits (Figure 1 – figure supplement 1).

**Figure 2. Enterovirus 3C$^{pro}$ cleaves human NLRP1 at the predicted site of mimicry and promotes pro-inflammatory cytokine release.** (A) Schematic of the domain structure of NLRP1, with predicted cleavage sites (triangles). FIMO-reported p-values and average

NetsurfP-reported coil probabilities (Figure 2 – figure supplement 1) are described at the predicted sites. (B) Percent conservation across 100 mammalian species at each position of each predicted 8mer cleavage site within human NLRP1. (C) Schematic of the human NLRP1 sequence used to assess enteroviral cleavage and activation. The predicted enteroviral cleavage site found in the linker region (127-GCTQGSER-134) is shown in red. Human NLRP1 WT-TEV contains an engineered TEV cleavage site between residues 93 and 94 (underlined green) in human NLRP1 WT. (D) Immunoblot depicting human NLRP1 cleavage by CVB3 $3C^{pro}$ and TEV protease. HEK293T cells were co-transfected using 100ng of the indicated Flag-tagged mCherry-NLRP1 fusion plasmid constructs with 250ng of the indicated protease construct and immunoblotted with the indicated antibodies. (E) Immunoblot depicting human NLRP1 cleavage at the indicated timepoints after infection with 250,000 PFU (MOI = ~1) CVB3. HEK293T cells were transfected using 100ng of either WT NLRP1 or NLRP1 G131P and infected 24-30 hours later. All samples were harvested 32 hours post-transfection and immunoblotted with the indicated antibodies (F) Immunoblot depicting human NLRP1 activation (maturation of IL-1β) by CVB3 $3C^{pro}$ and TEV protease. HEK293T cells were co-transfected using 100ng of the indicated protease, 50ng V5-IL-1β, 100ng CASP1, 5ng ASC, and 4ng of the indicated Myc-tagged NLRP1, and immunoblotted with the indicated antibodies. Appearance of the mature p17 band of IL-1β indicates successful assembly of the NLRP1 inflammasome and activation of CASP1. (G) Bioactive IL-1β in the culture supernatant was measured using HEK-Blue IL-1β reporter cells, which express secreted embryonic alkaline phosphatase (SEAP) in response to extracellular IL-1β. Supernatant from cells transfected as in (E) was added to HEK-Blue IL-1β reporter cells and SEAP levels in the culture supernatant from HEK-Blue IL-1β reporter cells were quantified by the QUANTI-Blue colorimetric substrate. Transfections were performed in triplicate and compared to the standard curve generated from concurrent treatment

of HEK-Blue IL-1β reporter cells with purified human IL-1β (Figure 2 – figure supplement 2). Data were analyzed using two-way ANOVA with Sidak's post-test. **** = $p < 0.0001$.

**Figure 3. Naturally occurring cleavage site variants alter NLRP1 susceptibility to enteroviral 3C$^{pro}$.** (A) Schematic of sites found to be evolving under positive selection (marked as *, from (Chavarria-Smith & Vance, 2013)) and human SNPs with at least 10 reported instances in the Genome Aggregation Database (GnomAD, (Karczewski et al., 2020)) (marked as !) within the linker region between the pyrin domain (PYD) and nucleotide binding domain (NBD) of NLRP1. The enteroviral 3C$^{pro}$ cleavage site between position 130 and 131 is indicated by a red triangle. (B) Phylogenetic tree depicting the enteroviral 3C$^{pro}$ cleavage site (red triangle) within NLRP1 across three clades of primates – hominoids, Old World monkeys (OWMs), and New World monkeys (NWMs). Mouse NLRP1B lacks any sequence that is alignable to this region of primate NLRP1 (see also Figure 3 – figure supplements 1 & 2). Amino acid differences to the human NLRP1 reference sequence are highlighted in red. Above the alignment is the enterovirus 3C$^{pro}$ sequence logo shown in Figure 1. (C) GnomAD-derived allele counts of each missense human SNP (by reference SNP #) within the 8mer of the determined enteroviral 3C$^{pro}$ cleavage site. (D-E) Immunoblot depicting CVB3 3C$^{pro}$ cleavage susceptibility of the indicated 8mer site variants introduced into human NLRP1 or full-length wild-type mouse NLRP1B (D) or the cleavage susceptibility of human NLRP1 Q130R, a naturally occurring human population variant (E). (F) Release of bioactive IL-1β into the culture supernatant as measured using HEK-Blue IL-1β reporter cells as in Figure 2G. Data were analyzed using two-way ANOVA with Sidak's post-test. **** = $p < 0.0001$.

**Figure 4. Diverse picornavirus 3C$^{pros}$ cleave and activate NLRP1 at independently evolved sites.** (A) Phylogenetic tree of 3C$^{pro}$ protein sequences for the indicated picornaviruses (Figure

4 – figure supplement 1 and 2). Shown next to the virus name is the sequence motif generated from the known sites of $3C^{pro}$ polyprotein cleavage in that specific virus. (B) Immunoblot depicting human NLRP1 cleavage by the indicated picornaviral $3C^{pro}$. Abbreviations are as in (A). Assays were performed as in Figure 2D. (left) Cleavage assays against WT NLRP1. (right) Human NLRP1 G131P mutant used in Figure 2. (C) Immunoblot depicting human NLRP1 cleavage at the indicated timepoints after infection with 250,000 PFU (MOI = ~1) CVB3 or EMCV. HEK293T cells were transfected using 100ng of either WT NLRP1 or NLRP1 G131P and infected 24-30 hours later. All samples were harvested 32 hours post-transfection and immunoblotted with the indicated antibodies (D) Release of bioactive IL-1β into the culture supernatant as measured using HEK-Blue IL-1β reporter cells as in Figure 2G. Data were analyzed using one-way ANOVA with Tukey's post-test. ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. (E) Immunoblot depicting mouse NLRP1B cleavage by numerous picornaviral $3C^{pro}$. Assays were performed as in Figure 2D.

**Figure 1 – figure supplement 1.** Motif optimization enhances capture of known human targets of enteroviral $3C^{pro}$. (A) As described in Figure 1B and 1C and Materials and Methods, the 8mer (P4-P4') $3C^{pro}$ polyprotein cleavage motif was initially generated from unique, concatenated 8mer cleavage sites across 796 enteroviral polyprotein sequences. To assess the capture capability of the motif on both virus and host targets, the motif was then used to conduct a low threshold (p-value = 0.1) FIMO (MEME Suite) search across training set of 2678 nonredundant enteroviral polyproteins from ViPR and 27 experimentally validated human targets of $3C^{pro}$ (Laitinen et al., 2016). In the graph, the X-axis represents a $\log_{10}$ of the p-value reported by FIMO as an indicator for the strength of the cleavage motif hit, or cleavage score. The left Y-axis depicts the number of uncalled 'true positives', or motif hits within the enteroviral polyprotein training set that overlap with the initial set of 8mer polyprotein cleavage sites used to generate

948 the motif (black). The right Y-axis depicts the number of called false positive sites, or any motif
949 hits that are not true positives, in the training set of enteroviral polyprotein sequences (gray).
950 (Above) Each line depicts a single, experimentally validated case of enteroviral $3C^{pro}$ cleavage
951 site within a human protein as reported in Laitinen et al, 2016 and is ordered along the x-axis by
952 its corresponding cleavage score. Vertical dotted lines are used to represent the decided
953 thresholds for comparison of capture capability. Capture of human targets at 95%, 99%, or
954 100% capture of true positives in the polyprotein dataset corresponds to capture of 4, 7, and 16
955 human hits. (B) Pseudo-counts to the position-specific scoring matrix of the motif shown in (A)
956 were adjusted by total information content where the two most information-dense positions P1
957 and P1' are assigned pseudocount = 0 and the least information-dense position P3
958 pseudocount = 1, and the remaining positions are assigned a pseudocount value relative to the
959 most information-dense position P1. This optimized motif is then used to FIMO search against
960 the same training set as described in (A). Capture of human targets at 95%, 99%, or 100%
961 capture of true positives in the polyprotein dataset corresponds to capture of 16, 23, and 24
962 human hits.
963
964 **Figure 2 – figure supplement 1.**
965 Tabular output of NetSurfP structural predictions for human NLRP1 describing the predicted
966 class (buried or exposed), relative surface accessibility, absolute surface accessibility,
967 probability for alpha-helix, probability for beta-strand, and coil probability for the amino acid
968 positions within each predicted enteroviral $3C^{pro}$ cleavage site.
969
970 **Figure 2 – figure supplement 2.**
971 Standard curve for Figure 2G. Purified human IL-1β was added in duplicate to the indicated final
972 concentration to HEK-Blue IL-1β reporter cells and SEAP activity was measured by increased

absorbance at $OD_{655}$. The indicated linear fit was used to calculate absolute concentrations of bioactive IL-1β from culture supernatants shown in Figure 2G. Note that supernatants from inflammasome-transfected cells was diluted 10-fold before addition to HEK-Blue IL-1β reporter cells to ensure that levels fell within the linear range of the indicated standard curve. Standard curves were generated in an identical manner for each panel of HEK-Blue data shown.

**Figure 3 – figure supplement 1.**

Mammalian NLRP1 phylogenomics and alignment of linker region. The indicated mammalian NLRP1 sequences were aligned and the region corresponding to residues 107-169 from human NLRP1 were extracted, which is anchored on both ends by well conserved proline and serine rich motifs. A consensus sequence generated from alignable sequences in this region is shown above the human sequence. The position of the CVB3 $3C^{pro}$ cleavage site in human NLRP1 is shown, flanked by four amino acids on both sides (P4->P4'). In other mammals, residues that differ from the human sequence are shown in red. Within the aligned region that corresponds to the CVB3 $3C^{pro}$ cleavage site, only simian primates have P4, P1 and P1' residues that would allow cleavage. The only other species that have a plausible cleavage site in this position are sheep and goats (P4 = Val, P1 = Gln, P1' = Ser), although those residues appear to have evolved independently at those positions. Two clades of species (the "mouse-related" clade of rodents and the microbat clade, marked as green) have NLRP1 protein sequences with N-terminal linkers that are unalignable to human NLRP1 in this region. Four additional clades (lagomorphs, megabats, cetaceans and felines, marked as grey) lack the NLRP1 gene altogether.

**Figure 3 – figure supplement 2.**

997 Schematic of the domain structure of mouse NLRP1B, with predicted cleavage sites shown as
998 colored triangles. FIMO-reported p-values are described at the predicted sites.
999
1000 **Figure 4 – figure supplement 1.**
1001 Alignment of $3C^{pros}$ used in this study. Sequences were aligned using MAFFT (Katoh &
1002 Standley, 2013) and used to generate the phylogenetic tree shown in Figure 4A. Asterisks
1003 indicate residues 100% conserved in all sequences. The position of the catalytic cysteine,
1004 analogous to C147 in CVB3 $3C^{pro}$, is highlighted in yellow.
1005
1006 **Figure 4 – figure supplement 2.**
1007 Table of pairwise percent sequence identity of $3C^{pros}$ used in this study as determined from the
1008 alignment shown in Figure 4 – figure supplement 1.
1009
1010 **Figure 4 – figure supplement 3.**
1011 Inhibition of NLRP1 activation by non-enteroviral $3C^{pro}$. HEK293T cells were transfected with
1012 inflammasome components as in Figure 2E using 100ng TEV protease or pQCXIP empty
1013 vector, but with the additional inclusion of a non-enterovirus $3C^{pro}$ or empty vector (100ng).
1014 Release of bioactive IL-1β into the culture supernatant was measured using HEK-Blue IL-1β
1015 reporter cells as in Figure 2G. Data were analyzed using one-way ANOVA with Tukey's post-
1016 test comparing all conditions containing TEV protease. **** = $p < 0.0001$.

## SUPPLEMENTARY FILES

**Supplementary file 1 – Related to Figure 1.**

**Training set of enteroviral polyproteins.** Accession IDs are listed for all polyproteins used to benchmark the motif search described in Figure 1D and Figure 1 – figure supplement 1. The 8mer cleavage sites and concatenated 8mer sequences are included.

**Supplementary file 2 – Related to Figure 1.**

**Enteroviral polyproteins with unique 8mer $3C^{pro}$ cleavage site concatenations.** Accession IDs are listed for all polyproteins used to create the search motif shown in Figure 1C and Figure 1 – figure supplement 1 and the enteroviral phylogenetic tree in Figure 1B. The 8mer cleavage sites and concatenated 8mer sequences are included.

**Supplementary file 3 – Related to Figure 1.**

**Un-optimized $3C^{pro}$ cleavage motif scores for true positive, false positive and human sites within the enteroviral polyprotein and human training sets.** FIMO-generated p-values and $\log_{10}$(p-value) represent the cleavage score at the provided matched sequence, where (A) is the un-optimized $3C^{pro}$ cleavage motif scores for true positive hits within enteroviral polyprotein dataset, (B) is the un-optimized $3C^{pro}$ cleavage motif scores for false positive hits within enteroviral polyprotein dataset where unique site matches are shown (26062), and (C) is the un-optimized $3C^{pro}$ cleavage motif scores for reported human cleavage sites from the Laitinen et al 2016 dataset.

**Supplementary file 4 – Related to Figure 1.**

**Optimized $3C^{pro}$ cleavage motif scores for true positive, false positive and human sites within the enteroviral polyprotein and human training sets.** FIMO-generated p-values and

1043 $log_{10}$(p-value) represent the cleavage score at the provided matched sequence, where (A) is the
1044 optimized $3C^{pro}$ cleavage motif scores for true positive hits within enteroviral polyprotein dataset,
1045 (B) is the optimized $3C^{pro}$ cleavage motif scores for false positive hits within enteroviral
1046 polyprotein dataset where unique site matches are shown (24437), and (C) is the optimized
1047 $3C^{pro}$ cleavage motif scores for reported human cleavage sites from the Laitinen et al 2016
1048 dataset.
1049
1050 **Supplementary file 5 – Related to Figure 2-4.**
1051 **List of primers and gBlocks used to clone protease- and NLRP1-encoded vectors.** Names
1052 and notes contain details on the restriction enzyme sites or point mutations encoded.
1053
1054 **Supplementary file 6 – Related to Figure 2-4.**
1055 **List of antibodies used for immunoblots.** Manufacturer and dilutions used are noted.
1056
1057 **Supplementary file 7 – Related to Figure 4-4.**
1058 **List of accession numbers used for sequence alignments**.
1059