A Unified Resource for Tracking Anti-CRISPR Names

Joseph Bondy-Denomy,1,* Alan R. Davidson,2,* Jennifer A. Doudna,3–9
Peter C. Fineran,10 Karen L. Maxwell,11 Sylvain Moineau,12 Xu Peng,13
Eric J. Sontheimer,14 and Blake Wiedenheft15

Dear editors

In the battle between CRISPR-Cas® prokaryotic immune systems and the elements that they target, a diverse array of “anti-CRISPR” proteins have evolved. These proteins appear to have arisen independently multiple times in evolution and function through diverse mechanisms to inhibit CRISPR-Cas immunity. For comprehensive reviews on anti-CRISPRs, we direct readers to recent publications.1,2

Due to the increasing interest in anti-CRISPRs, many new families of these proteins have been discovered in the past year or so. There are now 36 distinct families of anti-CRISPRs described in the literature that block seven subtypes of CRISPR-Cas systems.3–12 In 2015, a naming system for anti-CRISPR genes and proteins was introduced.6,13 To date, this system has been followed in all subsequent publications describing newly discovered anti-CRISPRs. However, as the rate of anti-CRISPR discovery will likely accelerate in the coming years, we feel that it would be advantageous to establish a database for the registration and tracking of anti-CRISPR names.

The primary goal of this database will be to prevent redundant names being used in publications, thus avoiding confusion in the literature. Anti-CRISPR proteins are named according to the subtype they inhibit and the order in which they were discovered—for example, AcrIF1 was the first anti-CRISPR protein identified to inhibit the type I-F system. The database (a Google document) can be found here: https://tinyurl.com/anti-CRISPR

We propose that this document be updated when researchers have had a manuscript accepted for publication in which new anti-CRISPRs are described. We suggest that the authors upload relevant data to the spreadsheet, including the name, CRISPR-Cas subtype inhibited, reference, and amino-acid sequence of the anti-CRISPR (Table 1). This spreadsheet may also be utilized by those preparing a manuscript for submission to ensure that they use anti-CRISPR names that are still available.

*Clustered Regularly Interspaced Short Palindromic Repeats.

Departments of 1Microbiology and Immunology, and 6Biochemistry and Biophysics, University of California, San Francisco, California; 2Departments of Biochemistry and Molecular Genetics, University of Toronto, Toronto, Ontario, Canada; Departments of 3Molecular and Cell Biology, and 4Chemistry, University of California, Berkeley, California; 5Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California; 6Gladstone Institutes, San Francisco, California; 7Howard Hughes Medical Institute and 8Innovative Genomics Institute, University of California, Berkeley, California; 9Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand; 10Department of Biochemistry, University of Toronto, Toronto, Canada; 11Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts; 12Department of Microbiology, University of Massachusetts Medical School, Worcester, Massachusetts; 13RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, Massachusetts; 14Department of Microbiology and Immunology, Montana State University, Bozeman, Montana.

*Corresponding authors.

Address correspondence to: Joseph Bondy-Denomy or Alan R. Davidson, E-mail: joseph.bondy-denomy@ucsf.edu or alan.davidson@utoronto.ca
To avoid the listing of many orthologues, we propose that the database only contain one entry per Acr, which will be considered the “type” Acr for that sequence family. In a case where a paper has investigated proteins that are homologous to an Acr protein, authors should utilize a subscript (e.g., AcrIF6_pae) to denote the species in which the anti-CRISPR is found. When multiple proteins from one species are investigated, we suggest a format of AcrIF6_pae-1, AcrIF6_pae-2, and so on. The established conventions for naming anti-CRISPR proteins and genes will be described as part of the database. We view this as an open repository for the field and as a complementary resource to a previously described anti-CRISPR database.14

Two of us (J.B.-D. and A.R.D.) were inspired to establish this database by the success of the CRISPR-Cas classification scheme in bringing order to the naming of Cas proteins.15,16 This work has been tremendously valuable for advancing the CRISPR-Cas field. We hope that our contribution to the anti-CRISPR field as presented here will provide a similar long-term benefit.

References


Table 1. A “Screenshot” from the Database, Depicting the Organization of Anti-CRISPR Entries

<table>
<thead>
<tr>
<th>Acr name</th>
<th>Type</th>
<th>Species of origin</th>
<th>Type of genomic element</th>
<th>Reference (First author, year, journal)</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcrIF1</td>
<td>I-F</td>
<td>Pseudomonas aeruginosa</td>
<td>Phage</td>
<td>Bondy-Denomy, 2013, Nature</td>
<td>MKFIKYLSTAHLNYMNIAYENGWS</td>
</tr>
<tr>
<td>AcrIF2</td>
<td>I-F</td>
<td>Pseudomonas aeruginosa</td>
<td>Phage</td>
<td>Bondy-Denomy, 2013, Nature</td>
<td>MIAQQHKDTVAACEAAEAIAIACKD</td>
</tr>
<tr>
<td>AcrIF3</td>
<td>I-F</td>
<td>Pseudomonas aeruginosa</td>
<td>Phage</td>
<td>Bondy-Denomy, 2013, Nature</td>
<td>MSSTISDRIIISRSVIEAARFIQSWE</td>
</tr>
</tbody>
</table>