

# Global AMP Survey database - GAS

1. Make it easy to rerun from any pc
2. We provided tables in tsv format and figures in ASCII  
(please read the Readme.txt file carefully)
3. Organize data sets in a database easy to update
4. Code the structure of webpage

# Home

Home

Browse

Tools

Downloads

Contact

About

Blah, blah, blah...

# Initial webpage

| Home | Browse      | Tools  | Downloads | Contact |
|------|-------------|--------|-----------|---------|
|      | AMPs        | Blastp |           |         |
|      | Families    | HMM    |           |         |
|      | Environment | Macrel |           |         |

# Browse-AMPs (Basic search)

- This is basically a large table with sort options, you can add all variables in the table we sent to you
- Search is a text wrapper
- Should there be options to select part of the table by using features intervals and etc.

# Browse-AMPs (Basic search)

|      |        |       |           |         |
|------|--------|-------|-----------|---------|
| Home | Browse | Tools | Downloads | Contact |
|------|--------|-------|-----------|---------|

Length

10-100aa

pl

11

Charge pH7

+3

Family

all

Search

AAKAKA

| AMP                     | Sequence        | pI   | Charge | Fam.                   |
|-------------------------|-----------------|------|--------|------------------------|
| <a href="#">GAS0001</a> | AALALALALAAKAKA | 11.1 | +3     | <a href="#">GAF035</a> |
| <a href="#">GAS0011</a> | AALALALALAAKAK  | 11.3 | +3     | <a href="#">GAF235</a> |
| <a href="#">GAS0021</a> | AAKAKACSDCK     | 11.4 | +3     | <a href="#">GAF075</a> |
| <a href="#">GAS1251</a> | AAKAKALAAKAK    | 11.1 | +3     | <a href="#">GAF055</a> |
|                         | .               |      |        |                        |
|                         | .               |      |        |                        |
|                         | .               |      |        |                        |

# Example for AMP browsing – dbAASP (<https://dbaasp.org/>)

Search Ranking Search

Saved Queries (?) <sup>NEW</sup>

Select an Option

[Add new](#) OR [Clear all](#)

## General Data

ID(s)

1 or 1,2,3, or 12-34



Name



Sequence



Options

☒ Full Sequence ☐ Part Of Sequence

Sequence Length

empty or 2- or -100 or 2-200



Complexity

All

Synthesis Type

All

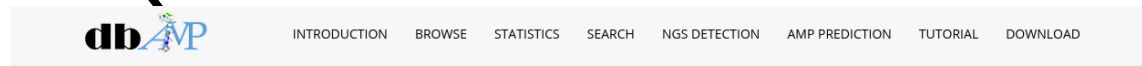
Uniprot ID

All

Found 20 / 14748 results

| ID | Name   | N terminus | Sequence   | C terminus | View                 |
|----|--|------------|--|------------|----------------------|
| 1  | Distinctin                                       |            | NLVSLGLEARKYLEQLHRKLKNKV<br>ENREVPFGFTALIKTLRCKII                                      |            | <a href="#">View</a> |
| 3  | Halocidin  |            | WLNALLHHGLNCAKGVLA<br>ALLHHGLNCAKGVLA  | AMD<br>AMD | <a href="#">View</a> |
| 4  | Khal   |            | KWLNALLHHGLNCAKGVLA<br>ALLHHGLNCAKGVLA   | AMD<br>AMD | <a href="#">View</a> |
| 6  | Enterocin X                                      |            | SNDSLWYVGQFMGKQANCITNHPV/KHMIIPGYCLSKILG<br>IAPIIVAGLGYLVKDAWDHSDQISGFKKGWNGRRK        |            | <a href="#">View</a> |
| 7  | EAFP1 + EAFP<br>2                                |            | XTCASRCRPPCNAGLCCSIYGYCGSGNAYCGAGNCRCQCRG<br>XTCASRCRPPCNAGLCCSIYGYCGSGAAYCGAGNCRCQCRG |            | <a href="#">View</a> |
| 8  |  | C16        | KVVvKVVvKvVK   | AMD        | <a href="#">View</a> |
| 9  | Lactococcin-Q<br>alpha + Lactoco<br>ccin-G alpha |            | SIWGDIGQVGKAAAYVVGKAMGNMSDVNQASRINRKKKH<br>GTWDDIGQIGIRVAYVVGKAMGNMSDVNQASRINRKKKH     |            | <a href="#">View</a> |
| 10 |  |            | LFIFFF   | AMD        | <a href="#">View</a> |
| 11 | Cathelicidin-1,<br>CATH-1, Fowlic<br>idin-1      |            | RVKRVWPLVIRTVIAGYNLYRAIKKK   |            | <a href="#">View</a> |
| 12 |  |            | RKRIHIGPGRAFYTT  |            | <a href="#">View</a> |
| 13 | RNal   | ACT        | RRXXRF   | AMD        | <a href="#">View</a> |
| 14 | Brevinin-2 relat<br>ed peptide B2R<br>P          |            | GIWDTIKSMGKVfAGKILQNL  | AMD        | <a href="#">View</a> |
| 15 | Histatin 5                                       |            | DSHAKRHHGYKRKFHEKHHSHRGY   |            | <a href="#">View</a> |
| 16 | Distinctin chain<br>1                            |            | ENREVPFGFTALIKTLRCKII  |            | <a href="#">View</a> |

# Example for AMP browsing – dbAMP (<http://140.138.77.240/~dbamp/index.php>)



## BROWSE

FORM FILTER

**Validation**

☐ Validated

☐ Predicted

**Against Activity**

☐ Antibacteria

☐ Anti-Gram(+)

☐ Anti-Gram(-)

☐ Anti-fungal

☐ Anti-viral

☐ Anti-parasitic

☐ Anti-HIV

☐ Wound Healing Activity

☐ Chemotactic

☐ Enzyme Inhibitor

☐ Malaria (e.g. P. falciparum)

☐ Cancer Cells

☐ Antimicrobial

☐ Anti-tumour

☐ Mammalian Cells

☐ Surface Immobilized

Totally : 12388 ( 1 / 496 pages)

|<12345>>|

| Browse all AMPs |        |  |          |
|-----------------|--------|--|----------|
| ID              | Length | Description  | Reviewed |
| dbAMP_00001     | 43     | <b>Source</b> : Helicoverpa armigera<br><b>Seq</b> : AAAAGSVWGA VNYTSDCNGECKRRGYKGGYCGSFANVNCWCET  | YES      |
| dbAMP_00002     | 18     | <b>Source</b> : Human Immunodeficiency virus 1<br><b>Seq</b> : AAANPGLLETSEGCRQIL  | NO       |
| dbAMP_00003     | 18     | <b>Source</b> : synthetic construct<br><b>Seq</b> : AAAPAATLEE HMTACQGV  | NO       |
| dbAMP_00004     | 97     | <b>Source</b> : Pleurotus citrinopileatus [Golden oyster mushroom]<br><b>Seq</b> : AACARFIDDFCDTLTPNIYRPRDNGQRCYAVNGHRCDFTVFNTNNGGNPIRASTPNC<br>KTVLRTAANRCP TGGGRKINPNAPFLAIDPNDGDCSTNF   | YES      |
| dbAMP_00005     | 40     | <b>Source</b> : Aurelia aurita [Moon Jellyfish]<br><b>Seq</b> : AACSDRAHGHCESFKSFGKDSGRNGVKLRANCKKTCGLC  | YES      |
| dbAMP_00006     | 60     | <b>Source</b> : Bos taurus [Bovine]<br><b>Seq</b> : AAEPDFYDSEEQMGPHQEADEKDRADQRVLT EEEKKELENLAAMDLELQKIAEF<br>SQR   | YES      |
| dbAMP_00007     | 255    | <b>Source</b> : Viola tricolor<br><b>Seq</b> : AAALPAAFASLET KDVITRAAYEKLVE SGAIQGITMTKTIISNPIL EALVAHFHKLGG<br>GTIFDCGESCF LGTCYTKGCS CGEWKLCYGTNSLPESNNEKAMVASLET KDVITRAAYEKLVN<br>SGAIQGITMTKTIISNPIL EALVAHFNRKLG GTIFDCGESCF LGTCYTKGCS CGEWKLCYGT<br>NSLPESNNEKAMVASLET KDVITRAAYENLVNSGAIQGITMTKTIISNPIL EALVSHFNRKLG<br>GGT | NO       |
| dbAMP_00008     | 102    | <b>Source</b> : Hybanthus floribundus subsp. floribundus<br><b>Seq</b> : AAALPAIATLEKDVITPETIQAVLKKTGSPLSNIMLEEDVINAMVKS KTVISKTVLDEALL<br>KHDGI GGIPCGFS CVFIPCISGVIGCSCKSKVCYRNSI IN   | NO       |

# AMP page

- It should contain all info for each AMP, such as features, cluster at 3 levels, location, metagenomes, environment
- Good example is provided, please give it a look and try to incorporate the tabs system
- The blue links are clickable and redirect the user to other webpages or also our own webpage
- Please find the metagenomes info in the SRA Database (<https://www.ncbi.nlm.nih.gov/sra/>) webpage



# AMP page (one for each)

[Home](#)[Browse](#)[Tools](#)[Downloads](#)[Contact](#)

## GAS001

|              |  |                |   |
|--------------|--|----------------|---|
| Sequence     | AALALALALAAKAKA  | Alternative ID | <a href="#">DRAMP001</a> , <a href="#">APD662</a> |
| Length       | 11.1   |                |   |
| Charge       | +3   |                |   |
| Family       | <a href="#">GAF035</a>   |                |   |
| Metagenomes  | <a href="#">SRR3856</a> , <a href="#">SRR6696</a> , <a href="#">SRR...</a> (connect with SRA database) |                |   |
| Environments | Pig guts, water, soil  |                |   |
| Genomes      | <a href="#">GAFSR.303049</a> (connect to NCBI)   |                |   |
| Species      | <a href="#">Bordethella cupris</a> (AMPs from same species)  |                |   |

[Features](#)

# Example for AMP info – dbAMP (<http://140.138.77.240/~dbamp/index.php>)

[INTRODUCTION](#)[BROWSE](#)[STATISTICS](#)[SEARCH](#)[NGS DETECTION](#)[AMP PREDICTION](#)[TUTORIAL](#)[DOWNLOAD](#)

## ANTIMICROBIAL PEPTIDE CARD

### General Information

[Basic](#)[Antimicrobial Potency](#)[Against Target Species](#)[Physical-chemical Properties](#)[External Links](#)

### Basic Information

|                       |  |
|-----------------------|--|
| dbAMP ID              | dbAMP_00001  |
| Name                  | Gallomycin   |
| Source                | Helicoverpa armigera   |
| Length                | 43   |
| Activity              | Nonrecorded  |
| Hemolytic Activity    | Unknown  |
| UniprotKB ID/AC       |  |
| Sequence              | AAAAGSVWGVNVTSDCNGECKRRGYKGGYCGSFANVNCWCET   |
| Description           | Gallomycin   |
| Taxonomy              | Eukaryota, Metazoa, Arthropoda, Insecta, Lepidoptera, Noctuidae, Helicoverpa, Helicoverpa armigera |
| Experimental Evidence | YES  |
| Target                |  |
| PubMed                |  |

# Protein family page

- Protein families are the clusters at the third level, we only have associated info to this level, therefore the levels I and II should be left as flat information
- We provided the alignment (aln), tree file (nwk), the three figure (txt), the HMM profile (hmm) and the hmmlogo (pdf); you should incorporate links to the download of these files to each family.
- Families were identified as GAF before the cluster number
- In families, there is the average and standard deviation for each feature, it should be taken in consideration.

# Protein family page (one for each)

|      |        |       |           |         |
|------|--------|-------|-----------|---------|
| Home | Browse | Tools | Downloads | Contact |
|------|--------|-------|-----------|---------|

Search

GAF001

| GAF001       |                       |                  |   |
|--------------|-----------------------|------------------|---|
| Sequences    | 36                    | Clustering level | 3 |
| Length avg.  | 11.1                  |                  |   |
| Charge avg.  | +3                    |                  |   |
| Environments | Pig guts, water, soil |                  |   |

HMM logo

Tree figure

Alignment

Tree\_file

HMM

# Example for protein families – PFAM (pfam.xfam.org/)



[HOME](#) | [SEARCH](#) | [BROWSE](#) | [FTP](#) | [HELP](#) | [ABOUT](#)



## Family: *Aminotran\_1\_2* (PF00155)

Loading page components (5 remaining)...

506 architectures

104216 sequences

1 interaction

9086 species

740 structures

### Summary

#### Domain organisation

#### Clan

#### Alignments

#### HMM logo

#### Trees

#### Curation & model

#### Species

#### Interactions

#### Structures

### Jump to...

enter ID/acc

Go

### Summary: Aminotransferase class I and II

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

No Wikipedia article: [Pfam](#) [InterPro](#)

This tab holds the annotation information that is stored in the Pfam database. As we move to using Wikipedia as our main source of annotation, the contents of this tab will be gradually replaced by the Wikipedia tab.

### Aminotransferase class I and II [Provide feedback](#)

No Pfam abstract.

#### Literature references

1. Nakai T, Okada K, Akutsu S, Miyahara I, Kawaguchi S, Kato R, Kuramitsu S, Hirotsu K; , Biochemistry 1999;38:2413-2424.: Structure of Thermus thermophilus HB8 aspartate aminotransferase and its complex with maleate. [PUBMED:10029535](#) [EPMC:10029535](#)
2. Ko TP, Wu SP, Yang WZ, Tsai H, Yuan HS; , Acta Crystallogr D Biol Crystallogr 1999;55:1474-1477.: Crystallization and preliminary crystallographic analysis of the Escherichia coli tyrosine aminotransferase. [PUBMED:10417420](#) [EPMC:10417420](#)

#### Internal database links

|                                     |  |
|-------------------------------------|--|
| SCOOP:                              | <a href="#">Alliinase_C</a> <a href="#">Aminotran_3</a> <a href="#">Aminotran_5</a> <a href="#">Aminotran_MocR</a> <a href="#">Beta_elim_lyase_Cys_Met_Meta_PP_DegT_DnrJ_EryC1</a> <a href="#">GDC-P_Met_gamma_lyase_OKR_DC_1_Pyridoxal_deC_SelA</a> <a href="#">SepSecS_SHMT</a>                                      |
| Similarity to PfamA using HHSearch: | <a href="#">Aminotran_3</a> <a href="#">Aminotran_5</a> <a href="#">Pyridoxal_deC_SHMT</a> <a href="#">DegT_DnrJ_EryC1</a> <a href="#">Cys_Met_Meta_PP_Beta_elim_lyase</a> <a href="#">OKR_DC_1</a> <a href="#">Alliinase_C</a> <a href="#">SepSecS</a> <a href="#">Met_gamma_lyase</a> <a href="#">Aminotran_MocR</a> |

#### External database links

|           |   |
|-----------|---|
| HOMSTRAD: | <a href="#">aat</a>                                 |
| PRINTS:   | <a href="#">PR00753</a> <a href="#">PR00799</a>     |
| PROSITE:  | <a href="#">PDOC00098</a> <a href="#">PDOC00518</a> |
| SCOP:     | <a href="#">1art</a>                                |

# Environment page (one for each)

[Home](#)[Browse](#)[Tools](#)[Downloads](#)[Contact](#)

Search

Soil

| Soil        |      |
|-------------|------|
| Sequences   | 36   |
| Length avg. | 11.1 |
| Charge avg. | +3   |

Families

Sequences

Compare with other environment

Genomes /  
Metagenomes list

Locations

# Tools - Blast webpage

| Home  | Browse                           | Tools   | Downloads                        | Contact                          |                                  |                                  |                                  |                                  |                                  |                                  |  |  |  |  |  |  |                         |  |  |  |  |  |                                   |  |  |  |  |  |
|---|----------------------------------|---|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|--|--|--|--|--|--|-------------------------|--|--|--|--|--|-----------------------------------|--|--|--|--|--|
| <p>Bastp search <input type="text" value="db"/></p> <div>Paste fasta sequences</div> <p><input type="button" value="Run"/> <input type="button" value="Upload file"/></p> |                                  | <p>Output</p> <table border="1"><thead><tr><th><input type="button" value="▼"/></th><th><input type="button" value="▼"/></th><th><input type="button" value="▼"/></th><th><input type="button" value="▼"/></th><th><input type="button" value="▼"/></th><th><input type="button" value="▼"/></th></tr></thead><tbody><tr><td colspan="6">Blast output (6 columns: qseqid, qtarget id, id, eval, score, qcovs)</td></tr><tr><td colspan="6">Fields with sort option</td></tr><tr><td colspan="6">Targets with link to the AMP card</td></tr></tbody></table> |                                  |                                  | <input type="button" value="▼"/> | <input type="button" value="▼"/> | <input type="button" value="▼"/> | <input type="button" value="▼"/> | <input type="button" value="▼"/> | <input type="button" value="▼"/> | Blast output (6 columns: qseqid, qtarget id, id, eval, score, qcovs) |  |  |  |  |  | Fields with sort option |  |  |  |  |  | Targets with link to the AMP card |  |  |  |  |  |
| <input type="button" value="▼"/>  | <input type="button" value="▼"/> | <input type="button" value="▼"/>  | <input type="button" value="▼"/> | <input type="button" value="▼"/> | <input type="button" value="▼"/> |                                  |                                  |                                  |                                  |                                  |  |  |  |  |  |  |                         |  |  |  |  |  |                                   |  |  |  |  |  |
| Blast output (6 columns: qseqid, qtarget id, id, eval, score, qcovs)  |                                  |   |                                  |                                  |                                  |                                  |                                  |                                  |                                  |                                  |  |  |  |  |  |  |                         |  |  |  |  |  |                                   |  |  |  |  |  |
| Fields with sort option   |                                  |   |                                  |                                  |                                  |                                  |                                  |                                  |                                  |                                  |  |  |  |  |  |  |                         |  |  |  |  |  |                                   |  |  |  |  |  |
| Targets with link to the AMP card   |                                  |   |                                  |                                  |                                  |                                  |                                  |                                  |                                  |                                  |  |  |  |  |  |  |                         |  |  |  |  |  |                                   |  |  |  |  |  |

# Tools - HMM webpage

|      |        |       |           |         |
|------|--------|-------|-----------|---------|
| Home | Browse | Tools | Downloads | Contact |
|------|--------|-------|-----------|---------|

## HMM search

Paste fasta sequences

RunUpload file

## Output

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| ▼ | ▼ | ▼ | ▼ | ▼ | ▼ | ▼ |
|---|---|---|---|---|---|---|

HMMScan output (6 columns:  
qseqid, qtarget id,  
eval, acc, qcovs)

Fields with sort option

Targets with link to the AMP card



# Example HMM search

(<https://pfam.xfam.org/search#tabview=tab1>)

## Search Pfam

Sequence

**Batch search**

Keyword

Domain architecture

Taxonomy

Jump to... 

enter ID/acc

[Go](#)

### Batch sequence search

Upload a FASTA-format file containing sequences to search for matching Pfam families using [the HMMER website](#).

Sequences file

[Browse...](#)

No file selected.

Cut-off

☐

Gathering threshold

☒

Use E-value

E-value

1.0

Email address

[Submit through the Hmmer website](#)

[Reset](#)

0  architectures

0  sequences

0  interactions

0  species

0  structures



**Pfam is part of the ELIXIR  
infrastructure**

Pfam is an Elixir service [Read more](#)

Comments or questions on the site? Send a mail to [pfam-help@ebi.ac.uk](mailto:pfam-help@ebi.ac.uk).

European Molecular Biology Laboratory

# Tools - Macrel webpage

Redirect to Macrel webpage or fuse in a similar way with the macrel webpage



## Antimicrobial activity prediction

You can paste peptides or contigs sequence in **FASTA** format into the field below:

input your sequence here

[Peptides example](#) || [Contigs example](#)

\* Data Type: ☐ Peptides ☐ Contigs (nucleotide)

Submit

Clear Form

For large inputs (>500KB or >1,000 sequences), please [download the tool](#) and run it locally.

Or submit a file in FASTA format:



Drag or click here to upload

(FASTA format, limited to 500KB)

Antimicrobial peptides (AMPs) are small proteins (10–100 amino acids) that can either cause lysis or generally interfere with cell growth. The rapid growth of antibiotic-resistant microorganisms in the last years became a world health problem. Thus, AMPs could represent a source of new antibiotic treatments. Recently, more and more genomic and metagenomic sequences have become publicly available. Macrel is pipeline that overcomes most of the problems associated to the prospection of AMPs in high throughput data sets analyzing diverse types of inputs (reads, contigs, protein sequences). The protein sequences predicted as AMPs by Macrel are, usually, specific and accurate, what can reduce the cost of experiments.

Macrel uses machine learning to select peptides with high probability of being an AMP. Furthermore, Macrel is also capable to perform the classification of AMPs into hemolytic and non-hemolytic peptides. This allows researchers to select the most interesting peptides for further testing *in vitro*.

Peptides submitted to the Macrel prediction should consist of 20 canonical amino acids and their length should range from 10 to 100 amino acids. Please avoid contigs containing non-canonical bases, such as N, R or Y.

This prediction system will output a table containing the predicted AMP sequence, its probability to be an AMP and the family it belongs, as well as the predictions

# Download HMM webpage

Home

Browse

Tools

Downloads


Contact

Fasta file containing AMPs to download

HMM profiles of families into a single  
DB for download


# Example of download page

(<http://140.138.77.240/~dbamp/download.php>)

INTRODUCTIONBROWSESTATISTICSSEARCHNGS DETECTIONAMP PREDICTIONTUTORIALDOWNLOAD



## DOWNLOAD

### Download All Antimicrobial peptides Sequence Data

| NAME     | INFO   | VERSION | FILE FORMAT | DATE                | ACTION  |
|----------|--|---------|-------------|---------------------|---|
| dbAMPseq | Release in 2019 - All published dbAMP Sequence data in EXCEL/FASTA format. | 2019-06 | XSLX, FASTA | 2019-06-15 15:23:24 |  |



[More dbAMPseq downloads...](#)

### Download Functional Activity Data

| NAME          | INFO   | VERSION | FILE FORMAT | DATE                | ACTION  |
|---------------|--|---------|-------------|---------------------|---|
| Antibacterial | Release in 2019 - All published functional activity Information for <b>Antibacterial</b> activity in EXCEL/FASTA format. | 2019-06 | XSLX, FASTA | 2019-06-15 13:07:52 |  |
| AntiGram_p    | Release in 2019 - All published functional activity Information for <b>Anti-Gram(+)</b> activity in EXCEL/FASTA format.  | 2019-06 | XSLX, FASTA | 2019-06-15 13:07:53 |  |

[More Activity downloads...](#)

### Catalog by Species

| NAME              | INFO  | VERSION | FILE FORMAT | DATE                | ACTION  |
|-------------------|---|---------|-------------|---------------------|---|
| dbAMPseq_bacteria | Release in 2019 - All published dbAMP Sequence data for <b>Bacterial</b> species in EXCEL format. | 2019-06 | XSLX        | 2019-06-25 12:30:46 |  |
| dbAMPseq_fungi    | Release in 2019 - All published dbAMP Sequence data for <b>Fungal</b> species in EXCEL format.    | 2019-06 | XSLX        | 2019-06-25 12:30:58 |  |

# Contact webpage

|      |        |       |           |         |
|------|--------|-------|-----------|---------|
| Home | Browse | Tools | Downloads | Contact |
|------|--------|-------|-----------|---------|

E-mail

Message

Send

Google docs form of satisfaction

Send

# Example of satisfaction form (<https://dbaasp.org/feedback>)

DBAASP<sup>V2.702</sup>

DATABASE OF ANTIMICROBIAL ACTIVITY  
AND STRUCTURE OF PEPTIDES

[Home](#) [Search](#) [Property Calculation](#) [Prediction<sup>NEW</sup>](#) [Statistics<sup>NEW</sup>](#) [About](#) [Help](#) [API](#) [Plugins](#) [Feedback](#)

Please, share your opinions and interests

Dear users

To get information from the users

Please point the page you have visited?

(For marking click on the circle)

☐ Search

☐ Ranking Search

☐ Prediction

☐ Property Calculation

☐ Statistics

Is the data of the peptide card comprehensive?

☐ Yes

☐ No

If "No", please state your opinion below about data should be added

Your answer