ESGEM-AMR

ESCMID

Agenda

1. First beta release of AMRrules

- Rules submitted by subgroups
- Code to annotate AMRfinderplus genotypes

2. Resource Updates

- Spec v0.6
- Open issues

3. Future Planning

- Timeline
- Papers
- Funding

4. *N. gonorrhoeae* group

- Defining rules for acquired mutations using AMRgen

https://github.com/interpretAMR/AMRrules/tree/main/rules

Core gene rules submitted and validated (v0.2-beta)

Available rules

Rule curation is a work in progress, under active development by the ESGEM-AMR Working Group.

Currently available rule sets are in the rules/ directory of this repository, named by organism. In this beta release they focus mainly on core genes and expected resistances, however acquired genes and mutations are included for some organisms already and will be added to others as the necessary data to define them accurately is accumulated and curated by the ESGEM-AMR working group.

- Acinetobacter baumannii
- Enterobacter
- Enterococcus faecalis
- Enterococcus faecium
- Escherichia coli
- Klebsiella pneumoniae
- Neisseria gonorrhoeae (acquired resistances, based on analysis of geno-pheno data)
- Pseudomonas aeruginosa
- Salmonella
- Staphylococcus aureus
- Yersinia

Note:
Bordetella and *Campylobacter* subgroups also submitted rules which are under validation.

Interpretation engine: installation

bit.ly/AMRrules

Beta version:

```
conda create -n amrrules_beta -c bioconda
python=3.12 pip

conda activate amrrules_beta

git clone https://github.com/interpretAMR/AMRrules
cd AMRrules

make dev
```

Current rules available **(v0.2-beta)**

- *Acinetobacter baumannii*
- *Enterobacter*
- *Enterococcus faecalis*
- *Enterococcus faecium*
- *Escherichia coli*
- *Klebsiella pneumoniae*
- *Pseudomonas aeruginosa*
- *Salmonella*
- *Staphylococcus aureus*
- *Yersinia*
- *Neisseria gonorrhoeae*
  (acquired resistances, based on analysis of geno-pheno data)

Interpretation engine: running

bit.ly/AMRrules

amrrules --input test_kleb_SGH10.tsv --output_prefix test_kleb_SGH10 --organism 's__Klebsiella pneumoniae'

Downloading Reference Gene Hierarchy from
https://ftp.ncbi.nlm.nih.gov/pathogen/Antimicrobial_resistance/AMRFinderPlus/database/latest/ReferenceGeneHierarchy.txt...
Downloaded and parsed 7685 rows from the Reference Gene Hierarchy.
4 hits matched a rule and 5 hits did not match a rule.
Output written to test_kleb_SGH10_interpreted.tsv.

amrrules --input test_data_amrfp_multiSpp.tsv --output_prefix test_multispp --organism_file test_data_sppCalls.tsv

Downloading Reference Gene Hierarchy from
https://ftp.ncbi.nlm.nih.gov/pathogen/Antimicrobial_resistance/AMRFinderPlus/database/latest/ReferenceGeneHierarchy.txt...
Downloaded and parsed 7685 rows from the Reference Gene Hierarchy.
All sample IDs from input file found in organism file. Proceeding with interpretation.
54 hits matched a rule and 106 hits did not match a rule.
Output written to test_multispp_interpreted.tsv.

Organism file example

SAMD00028844	s__Acinetobacter baumannii
SAMD00000649	s__Escherichia coli
SAMD00055728	s__Klebsiella pneumoniae
SAMD00002817	s__Staphylococcus aureus
SAMD00019033	s__Pseudomonas aeruginosa
SAMEA1572980	s__Neisseria gonorrhoeae
SAMD00499917	s__Yersinia enterocolitica

Matches "Name" in AMRFP input

Valid organism from --organism flag

Annotated AMRFP Report

bit.ly/AMRrules
(v0.2-beta)

NCBI **AMRfinderplus output**

AMRrules **genotype interpretation**

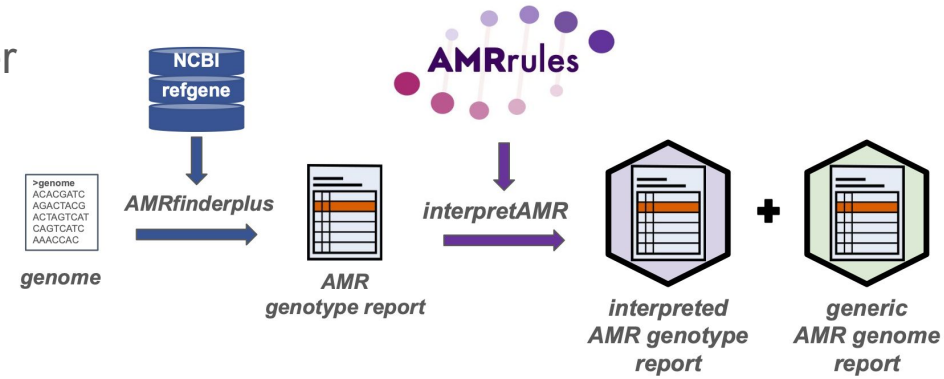
Name	Gene symbol	Class	Subclass	Hierarchy node	ruleID	context	drug	drug class	phenotype	clinical category	evidence grade	version	organism
SGH10	emrD	EFFLUX	EFFLUX	emrD	-	-	-	-	-	-	-	0.1.0	-
SGH10	oqxB19	PHENICOL/ QUINOLONE	PHENICOL/ QUINOLONE	oqxB19	KPN0003	core	ciprofloxacin	-	wildtype	S	moderate	0.1.0	s__Klebsiella pneumoniae
SGH10	oqxA	PHENICOL/ QUINOLONE	PHENICOL/ QUINOLONE	oqxA	KPN0002	core	ciprofloxacin	-	wildtype	S	moderate	0.1.0	s__Klebsiella pneumoniae
SGH10	blaSHV-11	BETA-LACTAM	BETA-LACTAM	blaSHV-11	KPN0001	core	-	penicillin beta-lactam	wildtype	R	strong	0.1.0	s__Klebsiella pneumoniae
SGH10	fosA	FOSFOMYCIN	FOSFOMYCIN	fosA5_fam	KPN0004	core	fosfomycin	-	wildtype	S	moderate	0.1.0	s__Klebsiella pneumoniae

- ‘minimal’ and ‘full’ annotation options (default is ‘minimal’)
- ‘full’ additionally contains:
  - breakpoint, breakpoint standard, evidence code, evidence limitations, PMID, rule curation note

Features planned for AMRrules v1.0

https://github.com/interpretAMR/AMRrules

- Ability to select AMRfp database version
- Ability to take hAMRonized AMRfp input
- Rule matching based on HMM accession
- Quality of life improvements
  - checking supported organisms
  - sensible warnings
  - limiting refgene download to once per run/install
  - conda installation
- **Summarised genome report**
  - requires drug and drug class logic



Example genome summary report

(v1.0 - in development)

name	drug	drug class	clinical category	phenotype	evidence grade	markers	ruleIDs	combo rules	organism
SAMN234	fosfomycin	-	S	wildtype	moderate	glpT_E448K (core)	ECO0082	-	s__Escherichia coli
SAMN234	-	penicillin beta-lactam	S	wildtype	strong	blaEC (core)	ECO0001	-	s__Escherichia coli
SAMN234	ciprofloxacin	fluoroquinolones	R	non wildtype	moderate	gyrA_S83L, parC_S80I	ECO0100, ECO0101	ECO0150	s__Escherichia coli
SAMN654	-	penicillin beta-lactam	S	wildtype	strong	blaSHV-11	KPN0001	-	s__Klebsiella pneumoniae
SAMN654	ciprofloxacin	fluoroquinolones	R	non wildtype	moderate	oqxA (core), oqxB19 (core), gyrA_S83L, qnrA1	KPN0003, KPN0002, KPN0035, KPN0047	KPN0132	s__Klebsiella pneumoniae

Agenda

1. **First beta release of AMRrules**
   - Rules submitted by subgroups
   - Code to annotate AMRfinderplus genotypes

2. **Resource Updates**
   - Spec v0.6
   - Open issues

3. **Future Planning**
   - Timeline
   - Papers
   - Funding

4. ***N. gonorrhoeae* group**
   - Defining rules for acquired mutations using AMRgen

Spec v0.6

Different breakpoints for multiple sites

E.g. cefuroxime for *E. coli* has breakpoints for iv and oral administration.

Cephalosporins[1]	MIC breakpoints (mg/L)			Disk content (µg)	Zone diameter breakpoints (mm)		
	S ≤	R >	ATU		S ≥	R <	ATU
Cefuroxime iv, *E. coli*, *Klebsiella* spp. (except *K. aerogenes*), *Raoultella* spp. and *P. mirabilis*	0.001	8		30	50	19	
Cefuroxime oral (uncomplicated UTI only), *E. coli*, *Klebsiella* spp. (except *K. aerogenes*), *Raoultella* spp. and *P. mirabilis*	8	8		30	19	19	

Current AMRrules guidance says to code 2 rules and indicate in the note, but need a more systematic way to record this besides note.

Propose: new column 'breakpoint condition'

Spec v0.6

Lack of evidence to support a species-specific interpretation of an acquired gene that has been found in a species

E.g. *dfr* allele found in Bordetella but with no matched AST from this genus

Propose: encode a rule with explicit evidence grade ('weak') and limitations ('lacks evidence for this species' if there is data for other members of the genus) or ('lacks evidence for this genus' if the only data is from outside the genus).

This is preferable to encoding no rule because:

- Without a rule, the choice of default interpretation is left to downstream reporting code
  - e.g. assuming that markers reported by AMRfp as 'class=Quinolones' implies NWT R for all quinolone drugs
- It explicitly highlights the uncertainty
  - Useful for AMRrules users to understand their genomes
  - Useful to the field to highlight where evidence gaps are

Spec v0.6

Added new ECO terms

ECO:0006404 **Experimentally evolved mutant phenotypic evidence.** A type of mutant phenotype evidence where mutations arising in organisms/cells over multiple generations, with or without exposure to some kind of selective pressure, are monitored or assayed.

STA0072	s__Staphylococcus aureus	rpoC	rpoC	WP_000567963.1			ARO:3003291	p.Phe632Ser	Protein variant detected

nonwildtype	R	MIC > 1	EUCAST v15.0 (2025-01-01)	16723576	ECO:0006404 experimentally evolved mutant phenotypic evidence	weak	unknown clinical relevance	This genotype is laboratory derived so it lacks clinical relevance.

Spec v0.6

Added new ECO terms

ECO:0000054 **Double mutant phenotypic evidence.** A type of mutant phenotype evidence resulting from an experiment typically constructed to determine if two different genes have an observable genetic interaction (functional connection) as the result of a mutation occurring in the alleles of the two genes of interest.

STA0053	s__Staphylococcus aureus	walK	walK	WP_000871610.1	ARO:3003794	p.Gly223Asp	Protein variant detected	acquired	vancomycin
STA0054	s__Staphylococcus aureus	graS		WP_001630781.1		p.Thr136Ile	Protein variant detected	acquired	vancomycin
STA0055	s__Staphylococcus aureus	STA0053 & STA0054					Combination	acquired	vancomycin

Spec v0.6

Specifying accessions for nucleotide-based genetic variants

e.g. 23S or promoter variants. as these are relative to a nucleotide sequence not protein sequence, there is no nodeID, HMM, or refseq/genbank protein accession.

refgene specifies these as 4 fields: `refseq nucleotide`, `refseq strand`, `refseq start`, `refseq stop` - e.g. https://www.ncbi.nlm.nih.gov/pathogens/refgene/#ampC_C-11T

In AMRfp output, it is identified by the '`accession of closest sequence`' field, like this:
"NZ_CP041538.1:1149245-1149489"
i.e., `refseq nucleotide`:`refseq start`-`refseq stop` (if forward strand)
or, `refseq nucleotide`:`refseq stop`-`refseq start` (if reverse strand)

Propose: add a nucleotide column amongst the set of possible NCBI unique identifiers, and adopt this format for encoding the nucleotide accession and coordinates to uniquely identify genes like 23S.

Spec v0.6

Taxonomy definitions

Currently we use GTDB as the reference standard, as this is genome-based.

However, there are problems for some of our target organisms, where species that we want to distinguish clinically are not differentiated in GTDB (which is based on consistent sequence divergence thresholds).

E.g.:

- Bordetella bronchiseptica is grouped under species Bordetella pertussis in GTDB
- Yersinia pseudotuberculosis is grouped under Yersinia pestis in GTDB

Propose: Add new field 'taxid' to indicate NCBI taxonomy ID (as this is stable as names change), and change definition of 'organism' field to be the 'current name' associated with the NCBI taxonomy ID.

Spec - open issues

Terms allowed in ‘context’ field?

Current options are ‘core’ or ‘acquired’

This field is intended to convey the context of the specific **variant** to which the rule applies, not just the **gene.**

This has been confusing, and many have encoded AMR-associated **mutations** in core genes as ‘core’.

- For presence of an acquired gene:
  - variation type = ‘Gene presence detected’
  - context = ‘acquired’
- For presence of a core gene:
  - variation type = ‘Gene presence detected’
  - context = ‘core’
- For an acquired mutation in a core gene:
  - variation type = ‘Protein variant detected’
  - context = ‘acquired’ … because this rule applies to the **mutation**, which is acquired

Spec - open issues

Terms allowed in ‘context’ field?

Current options are ‘core’ or ‘acquired’

This field is intended to convey the context of the specific **variant** to which the rule applies, not just the **gene.**
This has been confusing, and many have encoded AMR-associated **mutations** in core genes as ‘core’.

Propose: define ‘**gene context**’ as applying to the **gene** rather than the variant

- For presence of an acquired gene:
  - variation type = ‘Gene presence detected’
  - gene context = ‘acquired’
- For presence of a core gene:
  - variation type = ‘Gene presence detected’
  - gene context = ‘core’
- For an acquired mutation in a core gene:
  - variation type = ‘Protein variant detected’
  - gene context = ‘core’
- For a mutation in an acquired gene:
  - variation type = ‘Protein variant detected’
  - gene context = ‘acquired’

This also would address an issue raise by the Acinetobacter subgroup, who wanted to define two separate rules for

1) *Mutation in a chromosomal OXA gene*
2) *Same mutation in a mobile OXA gene*

With the proposed change, these would be separate rules, distinguished by ‘gene context’

Spec - open issues

How to code core efflux pumps that are reported by AMRfinderplus?

E.g. *amvA* in *Acinetobacter baumannii*

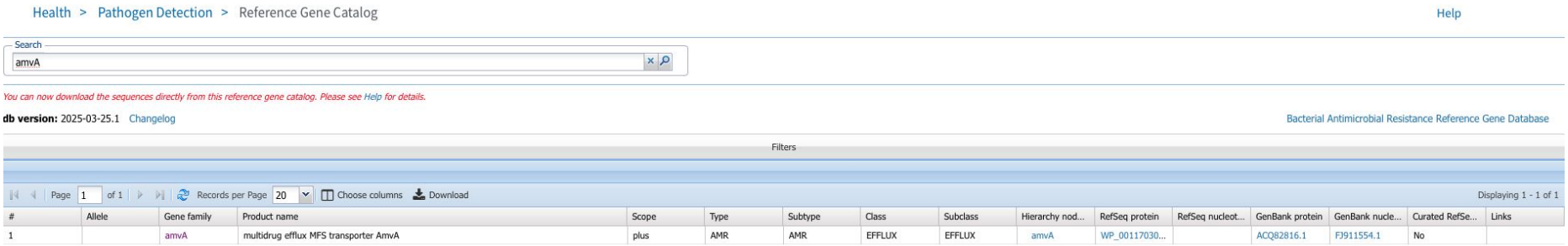
Do we need a rule?

✅They are reported in all genomes, and AMRrules could help users interpret this

❌If AMRfp annotates them as efflux (e.g. amvA), with no specific drug, so additional annotation is *not* needed?

✅CARD associated amvA with drug class ‘macrolide antibiotic’, so additional annotation *is* needed to correct that

Options: 1) encode with drug=blank, phenotype=wildtype, clinical category=blank (i.e. instead of S/I/R)
2) encode with class=efflux, as NCBI do?



Spec - open issues

Unknown AMR mechanisms for expected (intrinsic) resistance

E.g. ceftazidime in *Staphylococcus* spp.

Propose: Impose that all expected resistances should have a rule, even if the mechanism is unknown. This could be coded this as gene=unknown, context=core, interpretation = WT R.

Pros:

- Provide a comprehensive catalog of the known and unknown mechanisms behind expected resistance
- Provide a human-readable, curated explanation for why this is not yet known, which is useful for the community and also helps to explicitly highlight where knowledge gaps exist
- Ensuring expected resistances are encoded in AMRrules could assist with downstream phenotype reporting

Cons

- All AMRrules sets for all organisms would need to be reviewed to ensure they include rules for every expected resistance, even if there is no core gene to sensibly encode even a weak/null rule for (easier if we start now)
- Would need to ensure that the rules are 'complete' in this sense, with every new update of expected phenotypes
- One could argue that this kind of rule belongs in a downstream reporting logic, rather than in AMRrules

Spec - open issues

How to encode clinical category for a core gene when there is an expert rule to report as 'R', but it is not in the expected resistance list?

e.g. *aac(6')-Iaa* in *Salmonella enterica*

S. enterica isolates typically show MICs below the Enterobacterales breakpoints for these drugs (therefore they are not on the expected resistance list), but clinical failure with these agents has led to expert rules saying that *Salmonella* should always be reported R regardless of assay result.

Propose: define as 'wildtype R'... since the ultimate purpose of 'clinical category' is to predict treatment response NOT to predict in vitro assay result. (If the in vitro assay is not predicting clinical response well, there is probably a reason, such as this gene and/or the *mdsAB* efflux pump upregulating during infection and causing clinical failure. If we knew this, it would strengthen the rule and the call of R, but for now there is a lack of evidence on this).

- class = 'aminoglycosides' rather than specific agents, as that is how the expert rules are expressed
- breakpoint = 'not applicable' as the rules is specified for a class rather than a single agent
- breakpoint standard = 'EUCAST Salmonella Expert Rules vXX (year)'

Essentially we would be treating these expert rules the same as Expected resistances

Spec - open issues

Requests/issues from the subgroups

- Consider adding a field to explicitly indicate **gene family**
  - E.g. beta-lactamases (request from *Acinetobacter* group)
- How to encode rules for **inhibitors** or **disinfectants**?
  - Current spec allows values from CARD ARO ‘antibiotic’, this includes **disinfectants** already, so just need to make this explicit in the guidance
  - Allow **inhibitors** to entered in the ‘drug’ field?
    i. These are in CARD ARO as ‘adjuvant’ rather than antibitioc
    ii. Would simply require updating the spec to allow values from CARD ARO ‘adjuvant’ in addition to CARD ARO ‘antibiotic’
- Addition of fourth option in ‘clinical category’ field: 'S (inducible R)' ?
  - E.g. ‘inducible’ beta-lactamases in Enterobacter or ‘inducible’ clindamycin-resistance in Staph
  - This is departing from EUCAST/CLSI categorization
  - How to define? Is there a definition or standard for ‘inducible resistance’?

Agenda

1. **First beta release of AMRrules**
   - Rules submitted by subgroups
   - Code to annotate AMRfinderplus genotypes
2. **Resource Updates**
   - Spec v0.6
   - Open issues
3. **Future Planning**
   - Timeline
   - Papers
   - Funding
4. ***N. gonorrhoeae* group**
   - Defining rules for acquired mutations using AMRgen

Future Planning

Release of AMRrules v1.0 (July 2025)

- Features: Add genome-level summary output file
- Rules:
  - Add *Bordetella*, *Campylobacter*, others?
  - Still focused on core genes & expected R

Second release of AMRrules interpretation engine, v1.1 (Sept 2025)

- Rules: additional rules as available
- Features: accept hAMRonization formatted AMRfinderplus inputs

Quantitative approach to acquired resistance markers, v2.0 (Feb 2026)

- Rules: data-driven acquired gene rules for ESKAPEE + others
- Features: accept hAMRonization formatted CARD inputs

Timeline for quantitative rule generation

June: Quantitative group meetings (June 24), to discuss
- rule-setting criteria and code logic for AMRrulemakeR package
- what quantitative data fields to add to the AMRrules specification

July-Aug: Within organism subgroups
- discuss & test AMRrulemakeR package
- feedback issues and refinements needed for criteria / logic / specification

Sep-Dec: Generate and curate rules within organism subgroups

Jan-Feb: Aggregate rules into AMRrules v2.0 (Python package) release

Mar-Apr: Publish and present at Wellcome AMR Big Data & ESCMID Global

Papers

ESGEM-AMR core papers

1. Concept, specification, and rules for core genes/expected R
2. AMRgen package
3. Quantitative approach for acquired genes (& AMRrulesR package)
4. Analysis of mobile gene rules across organisms/groups

Individual subgroup papers? E.g. *N. gono*

Ideas? From the group - learnings from work so far:

- Gaps in knowledge about intrinsic resistance mechanisms
- Gaps in upstream AMR tools & databases

ESCMID Funding opportunities - reminder

ESCMID Research grants

- Study Group Grant: EUR30k, open 21 May–16 July
  - Organism subgroup?
  - Need to request ESGEM support by **21 June**
- Study Group Collaboration Grant: EUR180k, open 20 Aug–29 Oct
  - Generate high quality data to support quantitative rules for some spp?
  - Need to request ESGEM support by **20 September**

https://www.escmid.org/science-research/grants-awards/research-grants/

Agenda

1. First beta release of AMRrules

- Rules submitted by subgroups
- Code to annotate AMRfinderplus genotypes

2. Resource Updates

- Spec v0.6
- Open issues

3. Future Planning

- Timeline
- Papers
- Funding

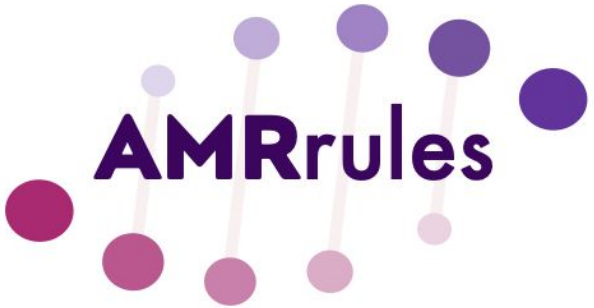
4. *N. gonorrhoeae* group

- Defining rules for acquired mutations using AMRgen

Questions? / Any other business?

ESGEM-AMR

ESCMID



https://github.com/interpretAMR/AMRrulesCuration