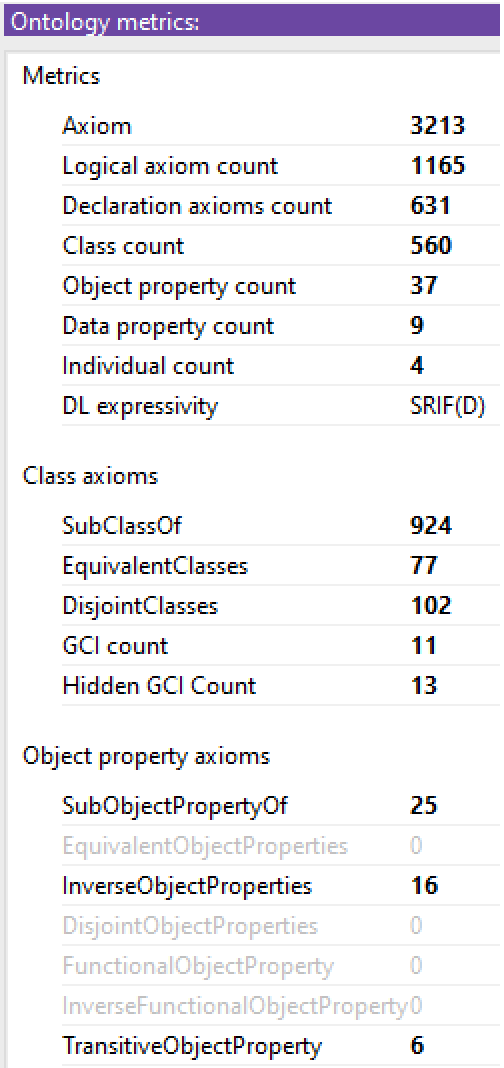
Gene regulation ontology (GRO)
Deconstruction

Stefan Schulz

Medical University of Graz, Austria

GRO - state before Malta Workshop

- Created 2008 within EU BootStrep
- Not maintained for a long time
- Available on BioPortal (OWL)
- URIs used by other projects (ISA, Kino, SYN)
- 2016 / 17 Revisited and revised for GREEKC
  - Alignment with upper-level ontology
  - Moderate content additions and modifications



Ontology metrics:

Metrics

Axiom	3213
Logical axiom count	1165
Declaration axioms count	631
Class count	560
Object property count	37
Data property count	9
Individual count	4
DL expressivity	SRIF(D)

Class axioms

SubClassOf	924
EquivalentClasses	77
DisjointClasses	102
GCI count	11
Hidden GCI Count	13

Object property axioms

SubObjectPropertyOf	25
EquivalentObjectProperties	0
InverseObjectProperties	16
DisjointObjectProperties	0
FunctionalObjectProperty	0
InverseFunctionalObjectProperty	0
TransitiveObjectProperty	6

GRO – decision @ Malta Workshop

- GRO development to be discontinued
- GRO content should be migrated to other ontologies
- More general: ontological content relevant for gene regulation should be covered by existing ontologies
- Rationale
  - Avoid proliferation of bio-ontology resources
  - Concentrate efforts on well-curated core ontologies

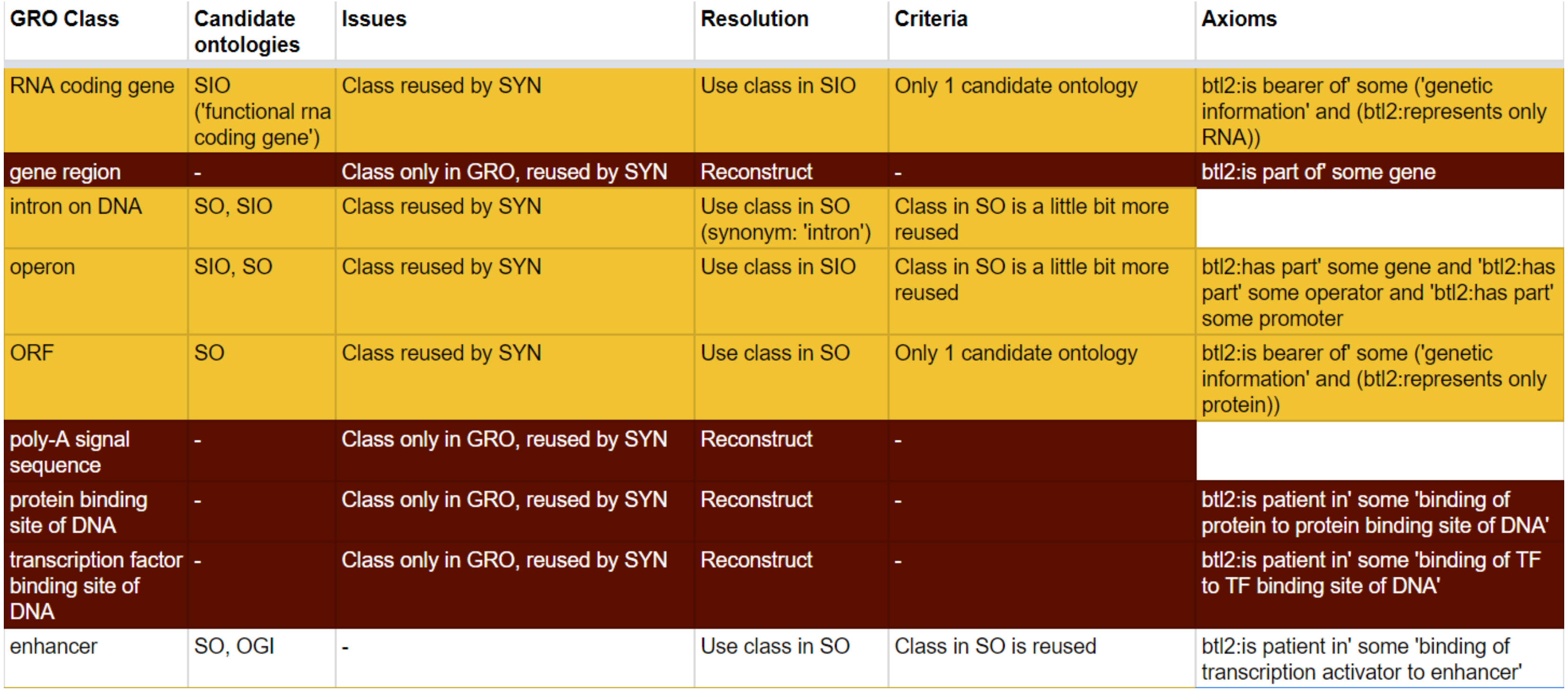
GRO: post-Malta agenda

- Analyse GRO content
- Identify gaps
- Identify overlaps
- Elaborate process for
  - GR ontological content acquisition
  - GRO content migration: identification of suitable target ontologies

GRO deconstruction

- Purposes:
  - Find new "home" for GRO classes
  - Align existing GRO terms with other ontologies
  - Select most "popular" or suited target ontologies
    - Preference: OBO ontologies, NCBI taxonomy
    - To avoid: Ontologies not obeying OWL semantics (NCIT), Ontologies not openly accessible (SNOMED CT)
  - Identify current use of GRO URIs
  - Assure persistence of currently used GRO URIs
  - Preserve axioms attached to GRO URIs

Alignment: current state https://goo.gl/64XaTt



GRO Class	Candidate ontologies	Issues	Resolution	Criteria	Axioms
RNA coding gene	SIO ('functional rna coding gene')	Class reused by SYN	Use class in SIO	Only 1 candidate ontology	btl2:is bearer of' some ('genetic information' and (btl2:represents only RNA))
gene region	-	Class only in GRO, reused by SYN	Reconstruct	-	btl2:is part of' some gene
intron on DNA	SO, SIO	Class reused by SYN	Use class in SO (synonym: 'intron')	Class in SO is a little bit more reused	
operon	SIO, SO	Class reused by SYN	Use class in SIO	Class in SO is a little bit more reused	btl2:has part' some gene and 'btl2:has part' some operator and 'btl2:has part' some promoter
ORF	SO	Class reused by SYN	Use class in SO	Only 1 candidate ontology	btl2:is bearer of' some ('genetic information' and (btl2:represents only protein))
poly-A signal sequence	-	Class only in GRO, reused by SYN	Reconstruct	-	
protein binding site of DNA	-	Class only in GRO, reused by SYN	Reconstruct	-	btl2:is patient in' some 'binding of protein to protein binding site of DNA'
transcription factor binding site of DNA	-	Class only in GRO, reused by SYN	Reconstruct	-	btl2:is patient in' some 'binding of TF to TF binding site of DNA'
enhancer	SO, OGI	-	Use class in SO	Class in SO is reused	btl2:is patient in' some 'binding of transcription activator to enhancer'

- 198 GRO classes checked, 74 not in other ontologies
- Most popular ontologies: SO, GO, ChEBI, FMA, PATO, BTO, SIO – Which should be preferred?

Next steps

- Complete alignment table
  - Currently done by Stefan and Jose
  - Needed: reviewers
- Devise migration strategy:
  - GRO URIs
  - GRO axioms
- Manage content inclusion requests
  - Contact curators of other ontologies
- GREEKG: "terminology observatory" role to assure high coverage of GR concepts