Clarify, simplify, and correct the top levels of the SCT Bacteria (organism) hierarchy.

Scope

Changes that might be made to the upper levels of the bacterial sub-hierarchy have the potential to affect all three user needs supported by organisms (laboratory reporting, public health and epidemiology and logical definitions within SCT).

Laboratory reporting

Content from the organism hierarchy is used for organism identification in laboratory reports. It is unlikely that changes in the upper hierarchy will affect this activity. There are examples of organisms that might be "reported" that are in the first one or two generations of the bacteria hierarchy (e.g. 372380008 | Campylobacter-like species) but these are rare. In the specific case cited, placement as a first generation descendent of Superkingdom Bacteria reflects known difficulty in defining the concept class.

Public health and epidemiology

Content from the organism hierarchy is used for organism identification in laboratory reports and many of these would be part of the public health "information stream." In the specific case of public health however, there is likely more interest in retrieval and aggregation based on the hierarchies and so it is more likely that Public health users might be affected by changes in the top one or two generations. We need to make an effort to reach out to this community for input concerning changes we might consider making.

Logical definitions in SCT

It would seem that the highest potential for adverse consequences associated with modifying the top levels of the bacteria sub-hierarchy would relate to unintended changes in the classification of infectious diseases. Appendix A, tables 1 and 2 enumerate uses of bacterial classes for definitions of disorders and procedures.

Classification of Bacteria (Systematics)

Dual classification schemes

Linnaean classes

There is no "official" bacterial taxonomy although there are three primary references that are often cited as "official" sources for the organization and classification of bacteria:

- The List of Prokaryotic Names with Standing in the Nomenclature (LPSN)
- Bergey's Manual of Determinative Bacteriology
- Bergey's Manual of Systematic Bacteriology

Each of these references is very clear and deliberate in saying that when they report bacterial classification(s) they are reporting what they believe is a consensus but each declares that any taxonomy stated or implied is not official. Bacterial taxonomy (really all of living organism taxonomy) is always in flux. Opinions of competing experts, often speaking to a different BASIS for the classification of various organisms can be discerned in all reference sources on the subject.

There are 7 primary Linnaean taxonomic classes, namely Domain, Kingdom, Phylum, Class, Order, Family, Genus, Species. Intercalated between these one can also find a number of intermediate classes such as Superkingdom, Infraorder, Subclass and Subspecies. Intermediate classes may or may not be present in the lineage of any particular bacterial species. In fact intermediate classes are often invoked specifically to represent areas of taxonomy that are controversial or unsettled. As an example, the (current) top level of the bacteria hierarchy 409822003 | Superkingdom Bacteria, is now considered to be "Domain Bacteria".

"Phenotypic" classes

Classes based on phenotypic characteristics of bacteria (gram negative, anaerobic, bacillus, etc.) are used for laboratory findings such as 404510009 | Large gram-negative rods (finding) and for defining clinical infections such as 449082003 | Sepsis caused by Gram negative bacteria (disorder). It remains to be determined how important these phenotypic classes are for retrieval and aggregation by users of SCT.

Issues for phenotypic classes on SCT:

1. 41146007 | Bacterium
   a. FSN does not distinguish from 409822003 | Superkingdom Bacteria
   b. May have been (likely) created as a "navigational" class for more granular phenotypic categories
   c. Its direct use as value for SCT definitions should all be assigned to 409822003 | Superkingdom Bacteria
   d. 5 of 32 "first descendant" phenotypic classes are assigned to 409822003 | Superkingdom Bacteria

2. 409812001 | Actinobacteria
   a. FSN does not distinguish from 409822003 | Superkingdom Bacteria
   b. May have been (likely) created as a "navigational" class for more granular phenotypic categories
   c. Its direct use as value for SCT definitions should all be assigned to 409822003 | Superkingdom Bacteria
   d. 2 of 12 "first descendant" phenotypic classes are assigned to 409822003 | Superkingdom Bacteria
Inconsistencies in naming

2. Inconsistencies in naming
   a. 115173005 | Facultatively anaerobic gram-negative rod is a subtype of 87172008 | Gram-negative bacillus which is a subtype of 416495007 | Form-bacillus
   b. 115173005 | Facultatively anaerobic gram-negative rod is a subtype of 414185006 | Facultative anaerobic bacteria

3. Completeness
   a. 557 SCT organism classes are descendants of 409822003 | Superkingdom Bacteria but are not descendants of 41146007 | Bacterium.
      i. although many of these may not be medically significant some of them certainly are (e.g. Treponema pallidum)
      ii. a few are phenotypic classes (e.g. 415625001 | Strict aerobic bacteria)
   b. Are all important / useful phenotypic classes present?
   c. Are subtype relationships between phenotypic classes complete?
   d. Are Linnaean classes assigned to phenotype classes where known, appropriate and necessary?

4. Accuracy
   a. Are phenotype class assignments accurate?
      i. Review as of April 2017 suggests that they (generally) are.

Intercaletion of the schemes

Historically (at least) the phenotypic characteristics of bacteria are important tools for their identification. This creates the (usually correct) direct subtype relationships between the phenotypic classes and the Linnaean classes seen in SCT. For example, all subtypes of 409852006 | Phylum Proteobacteria (organism) are gram negative and 81325006 | Gram-negative bacterium (organism) is its only immediate supertype. These two schemes are interwoven throughout the sub-hierarchy descending from 409822003 | Superkingdom Bacteria (organism). Not all Linnaean classes have supertypes in the phenotype hierarchy. This may be because particular bacteria (e.g. 426497002 | Phylum Acidobacteria) have not been characterized sufficiently to place them according to their phenotype. Further, there is no known medical importance to these organisms (they are found in soil and do not cause infections) there is no urgency to connect them to the phenotypic hierarchy.

For discussion:

1. Consider editorial policies for Linnaean hierarchy (which are and generally should be established for the entire organism hierarchy).
   a. Are there exceptions to the general policy that are required for correct and useful classification of bacteria?
2. Develop editorial policies for the phenotypic hierarchy.
   a. Bacteria are among a very few for which an extensive phenotypic hierarchy has been developed.
   b. Is the phenotypic hierarchy of the bacteria a special case, a general case, both?
3. Develop editorial policies for the structure and interconnections between the two hierarchies.
   a. Is it necessary to replace “Bacterium” with something more understandable and useful?
   b. Is it necessary to review the bacteria and correct (all) assignments of Linnaean class to (known) phenotype?

- Appendix A. Descendants of Superkingdom Bacteria and Bacterium