

# Report on status of mouse phenotype ontologies

George Gkoutos  
University of Cambridge

Corfu  
September 2007

# Phenotype And Trait Ontology (PATO)

## An ontology of phenotypic qualities

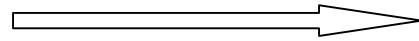
- Phenotypes may be described in many different dimensions, e.g.
  - the biochemical ('alcohol dehydrogenase null')
  - the cellular ('cell division arrested at metaphase'),
  - the anatomical ('eye absent')
  - the behavioral ('hyperactive').
  - etc.
- In whatever dimension and granularity, however, there is a commonality and the great majority of, if not all, phenotypic descriptions can be decomposed into two parts
  - An entity that is affected. This entity may be an enzyme, an anatomical structure or a complex biological process.
  - The qualities of that entity.

# Features of Qualities

- Qualities are the basic entities that we can perceive and/or measure:
  - colors, sizes, masses, lengths etc.
- Qualities inhere to entities: every entity comes with certain qualities, which exist as long as the entity exist.
- Qualities belong in a finite set of quality types (i.e. color, size etc) and inhere in specific individuals. No two individuals can have the same quality, and each quality is specifically constantly dependent on the entity it inheres in.

# Phenotypic Character

Phenotypic Character



entity + quality

(mouse body weight)

(mouse anatomy: body + PATO: weight)

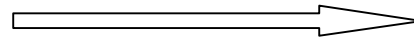
(eye colour)

(Drosophila anatomy: eye + PATO: colour)

(glucose concentration)

(ChEBI: glucose + PATO: concentration)

increased size hepatocellular carcinoma



hepatocellular carcinoma (MPATH:357) has\_quality increased size (PATO:0000586)

# Representation of Phenotypic data

Assay (eg. Histopathology, blood chemistry) *{constrained\_by}*

MGED (environmental & genetic)- *{of type}* - **conditions**

{ Entity (eg. MA) - Quality (PATO)  
Entity (eg.MPATH) - Quality (PATO)  
Entity (eg. Cell) - Quality (PATO)

# Annotation Issues

- EQ (eye red) is not enough
  - Fine for simple subsets
- Extensions:
  - time
  - quantification
    - EQuant={some,all,most,<percentage>,<count>}
    - E=thoracic bristle, Q=long, EQuant=80%
      - 80% of the thoracic bristles *in this one individual fly*
  - relational qualities
    - sensitivity
  - comparison
  - modifiers (low, medium, high etc)
  - measurements
  - ...

# Measurements

- Ontologies provide *qualitative* partitions on the kinds of entities we find in nature
- We may also want to record *quantitative* information
  - Comes from *measurements* of qualities
  - The measurement is not the phenotype
    - Phenotypes exist independently of our measurements of them

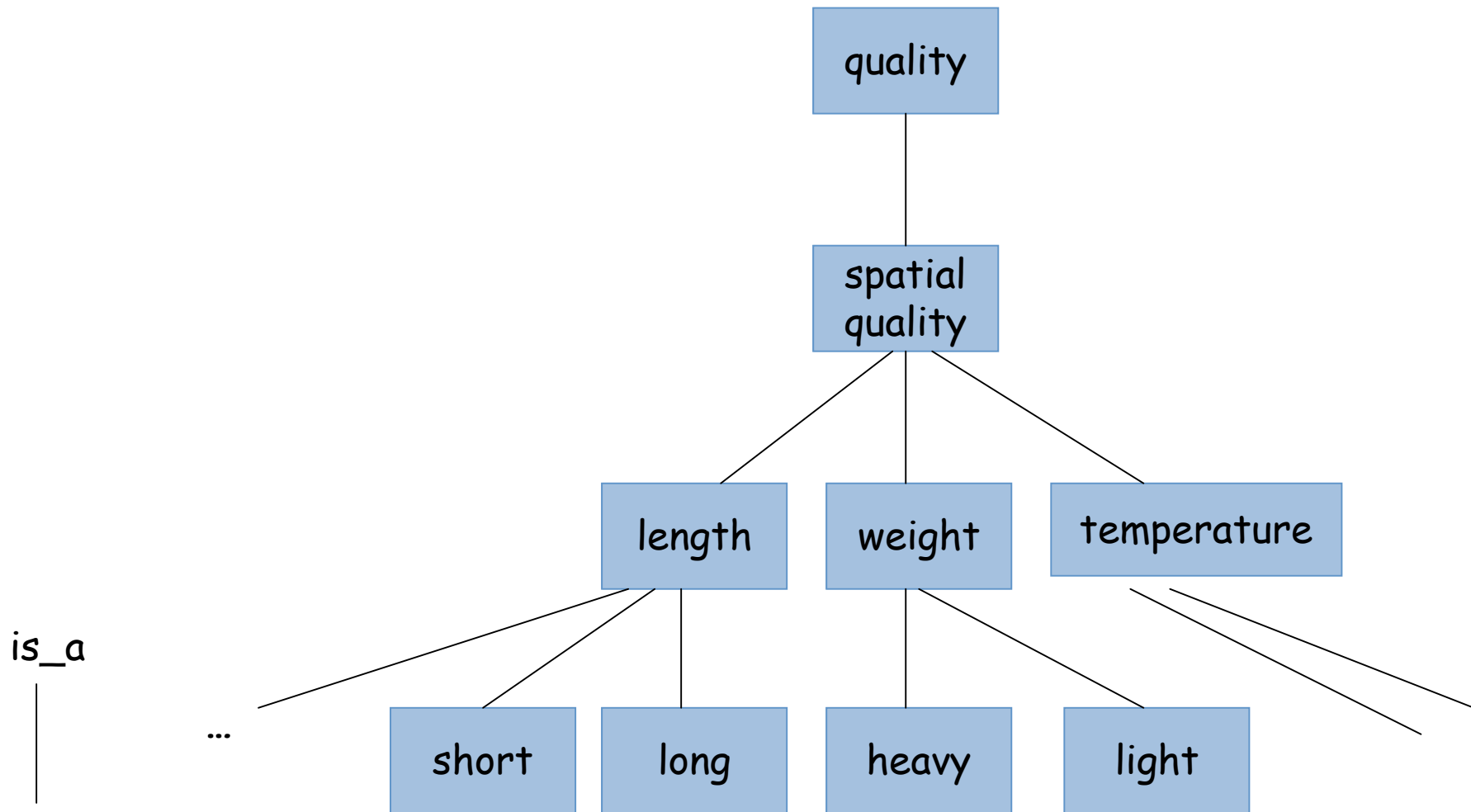
# Scalar qualities

The tail of my mouse is 2.1 cm

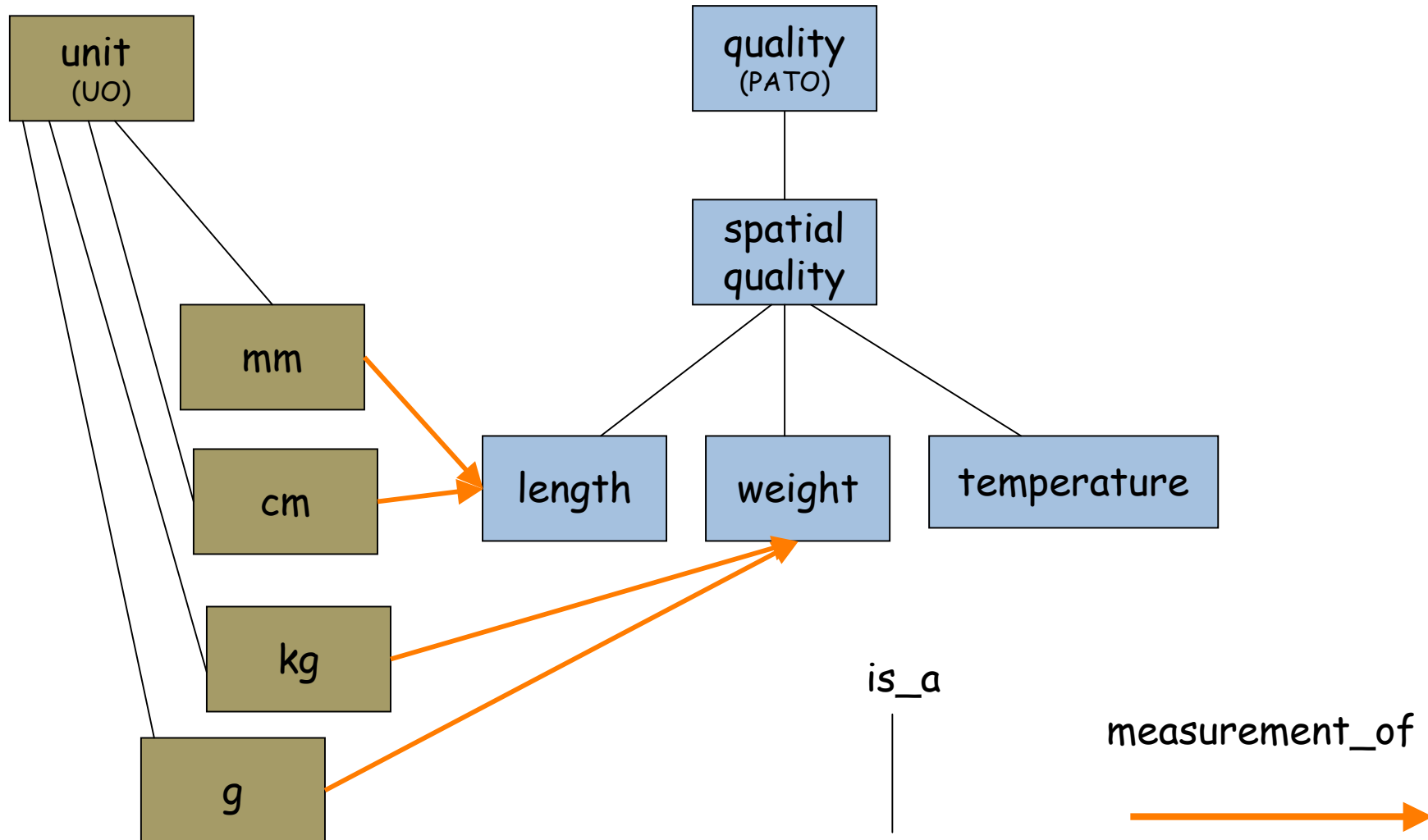
- A scalar quality can be partitioned on a linear scale
- Scalar qualities can be measured
- Measurements involve units



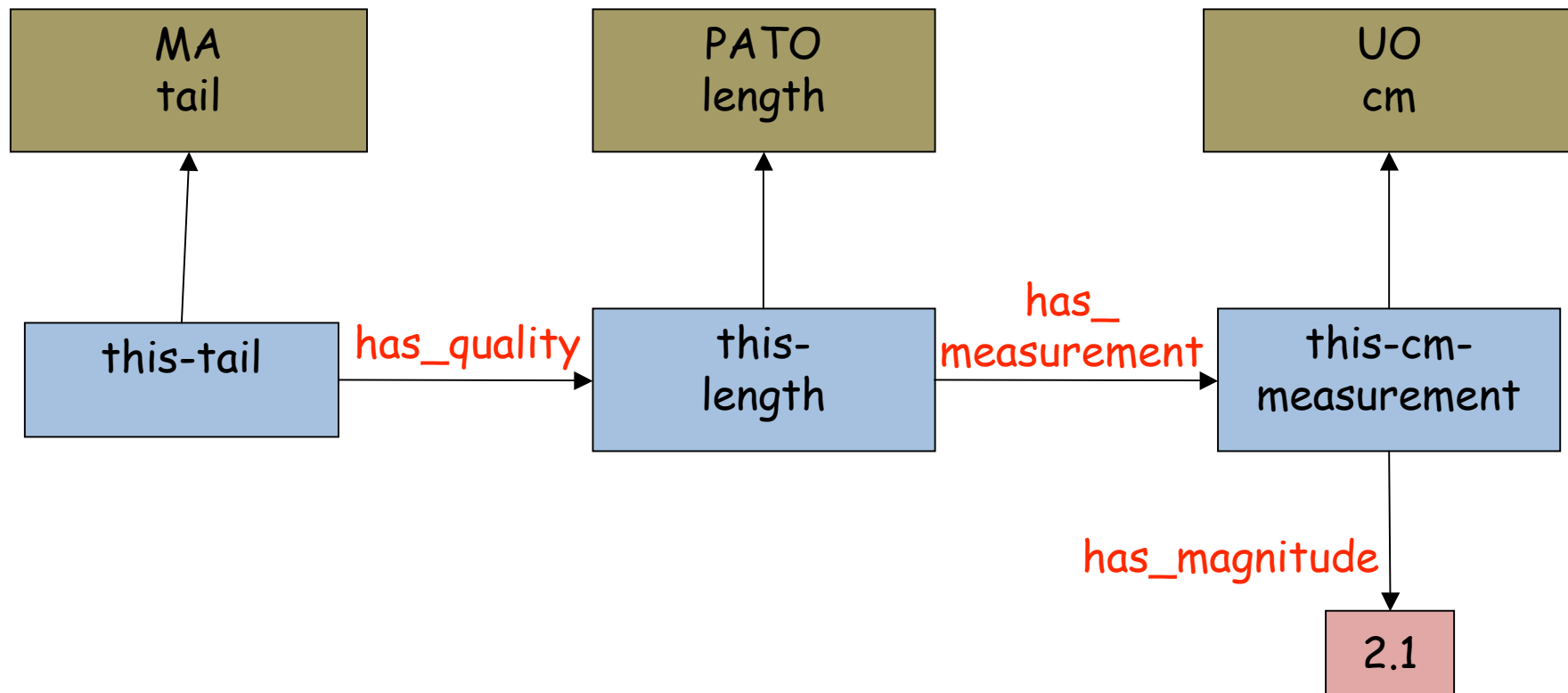
# PATO & scalar qualities



# Representation of measurements



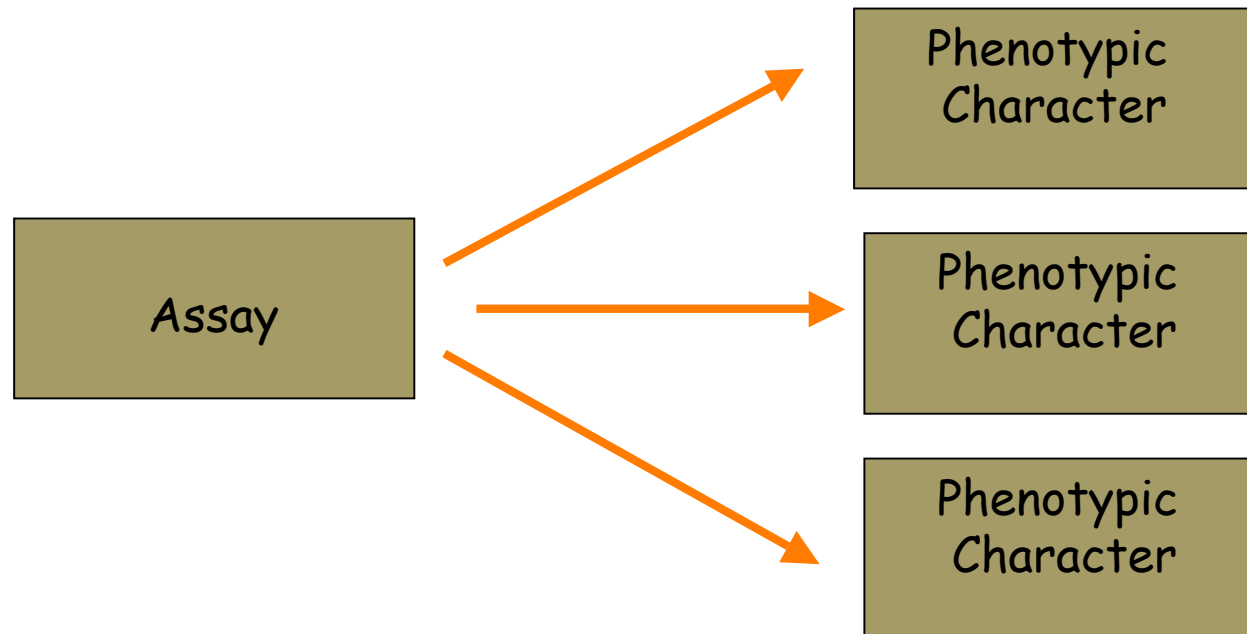
# Measurement instances in OWL



# Assay Controlled Vocabulary

- Abnormality
- Relative\_to
- Ranges of values
- Allows the schema to be dynamic

- Definition of qualities and their relations
- Explicit differences (between laboratories)
- Allows labs around the world to “plug-in” their assays to the schema



# PATO Annotations

- Descriptions can be pre- or post- composed
- Post-coordinated
  - NCBO fly-fish-human disease gene annotations
  - BIRN image annotation (neurodegenerative disease)
  - NESCent / AToL - evolutionary character matrices
  - ...
- Pre-coordinated
  - MGI Mouse genotype-phenotype annotation (Mammalian Phenotype)
  - Gramene trait annotation (Plant trait ontology)
  - ...

# Reconciling pre and post composed annotations

- Retrospective PATO definitions of pre-coordinated terms in phenotype ontology
  - Example
    - "potassium sensitivity" = "sensitivity"<sub>PATO</sub> towards "potassium"<sub>ChEBI</sub> (*inheres\_in* "plant"<sub>PSO</sub>)
    - "decreased bone mass" < "decreased mass"<sub>PATO</sub> *inheres\_in* "bone"<sub>MA</sub>
- Current progress
  - Mammalian Phenotype: 4083 / 5579
  - Plant trait : 340 / 765
  - Worm phenotype: 422 / 1520

# Decomposition of MPO

- 5579 terms - 4083 of which are decomposed.
  - First pass by OBOL (automatic) around 400 terms
  - Second pass by hand
- Ontologies used
  - GO
  - MA
  - ChEBI
  - MPATH
  - Cell
  - EMAP

- [-] abnormal adipose tissue a
  - [+] abnormal brown adip
  - [+] abnormal percent bo
  - [+] abnormal white adip
  - [-] decreased adipose t
  - [-] increased adipose ti
  - [+] abnormal adipose tissue d
  - [+] abnormal brown adipose t
  - [+] abnormal fat pad
  - [+] abnormal white adipose ti
  - [+] abnormal adipose tissue physiol
- [+] behavior/neurological phenotype
- [+] cardiovascular system phenotype
- [+] cellular phenotype
  - [-] abnormal cell content/ morphol
  - [-] abnormal cell mass
    - [-] decreased cell mass
    - [-] increased cell mass
  - [+] abnormal lysosome morph
  - [-] abnormal mitochondrial m
  - [+] abnormal nucleus count
  - [+] abnormal nucleus morpho
  - [+] abnormal plasma membra
  - [-] abnormal cell migration
  - [+] abnormal cell number
  - [+] abnormal cell physiology
- [+] craniofacial phenotype
- [+] digestive/alimentary phenotype
- [+] embryogenesis phenotype
- [+] endocrine/exocrine gland phenotype
- [+] growth/size phenotype
  - [-] abnormal postnatal growth/weig
  - [-] abnormal body size
    - [+] abnormal body heig
    - [+] abnormal body leng
    - [-] abnormal body weig
    - [+] decreased bod
    - [+] increased bod
    - [-] decreased body size
    - [-] increased body size
  - [+] abnormal chest morpholo
  - [+] abnormal lean body mass
  - [+] abnormal postnatal growth
  - [-] distended abdomen
  - [-] heterotaxia
  - [+] left-sided isomerism
  - [+] right-sided isomerism
  - [-] situs ambiguus
  - [-] situs inversus
- [+] abnormal prenatal growth/weight
- [+] hearing/vestibular/ear phenotype
- [+] hematopoietic system phenotype
- [+] homeostasis/metabolism phenotype
- [+] immune system phenotype
- [+] lethality-embryonic/perinatal
- [+] lethality-postnatal
- [+] life span-post-weaning/aging

intersection\_of: PATO:0000573 ! increased length  
 intersection\_of: inheres\_in MA:0002405 ! adult mouse

[Term]  
 id: MP:0001258 ! decreased body length  
 intersection\_of: PATO:0000574 ! decreased length  
 intersection\_of: inheres\_in MA:0000004 ! trunk

[Term]  
 id: MP:0001259 ! abnormal body weight  
 intersection\_of: PATO:0000128 ! weight  
 intersection\_of: qualifier PATO:0000460 ! abnormal  
 intersection\_of: inheres\_in MA:0002405 ! adult mouse

[Term]  
 id: MP:0001260 ! increased body weight  
 intersection\_of: PATO:0000582 ! increased weight  
 intersection\_of: inheres\_in MA:0002405 ! adult mouse

[Term]  
 id: MP:0001262 ! decreased body weight  
 intersection\_of: PATO:0000583 ! decreased weight  
 intersection\_of: inheres\_in MA:0002405 ! adult mouse

[Term]  
 id: MP:0001264 ! increased body size  
 intersection\_of: PATO:0000586 ! increased size  
 intersection\_of: inheres\_in MA:0000004 ! trunk

[Term]  
 id: MP:0001267 ! enlarged chest  
 intersection\_of: PATO:0000586 ! increased size  
 intersection\_of: inheres\_in MA:0000031 ! chest

[Term]  
 id: MP:0001270 ! distended abdomen  
 intersection\_of: PATO:0001602 ! distended  
 intersection\_of: inheres\_in MA:0000029 ! abdomen

[Term]  
 id: MP:0001274 ! curly vibrissae  
 intersection\_of: PATO:0000405 ! curled  
 intersection\_of: inheres\_in MA:0000163 ! vibrissa

[Term]



# Some Issues

- Some structural and definition issues
- Core ontologies term requests
  - GO (~100)
  - MA (~140)
  - EMAP (~60)
  - ChEBI (~10)
  - Cell (~30)

- Phenotype and Units Mailing list
  - PATO  
<http://www.obofoundry.org/cgi-bin/detail.cgi?id=quality>
  - UO  
[http://sourceforge.net/mailarchive/forum.php?forum\\_name=obo-unit](http://sourceforge.net/mailarchive/forum.php?forum_name=obo-unit)
- PATO and UO Term request tracker
  - PATO  
[http://sourceforge.net/tracker/?group\\_id=76834&atid=595654](http://sourceforge.net/tracker/?group_id=76834&atid=595654)
  - UO  
[http://sourceforge.net/tracker/?group\\_id=76834&atid=595654](http://sourceforge.net/tracker/?group_id=76834&atid=595654)
- Phenote Tool
  - <http://www.phenote.org/>



# Representation of disease data

*Pathology is an attribute of a tissue as it represents a tissue response to an underlying lesion*

**Qualifier** (histological type, grade, etc. eg metastatic)-**Entity**  
(MPATH)- {has} - **Quality** (PATO)

Assay

|

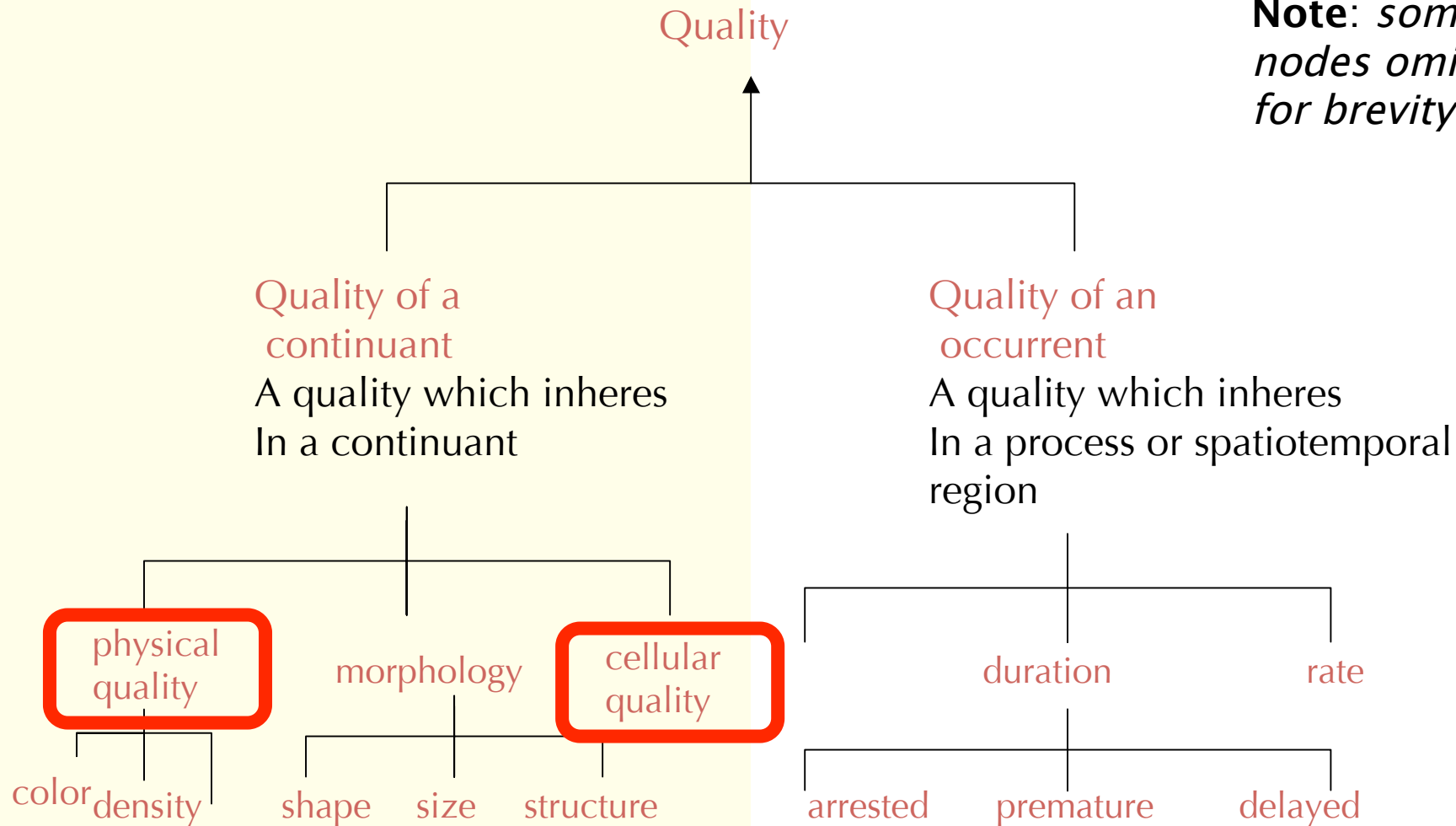
**Histopathological Observation** -  
{constrainedby}

|

(environmental & genetic)- {oftype} - **conditions**

# PATO: Top level division

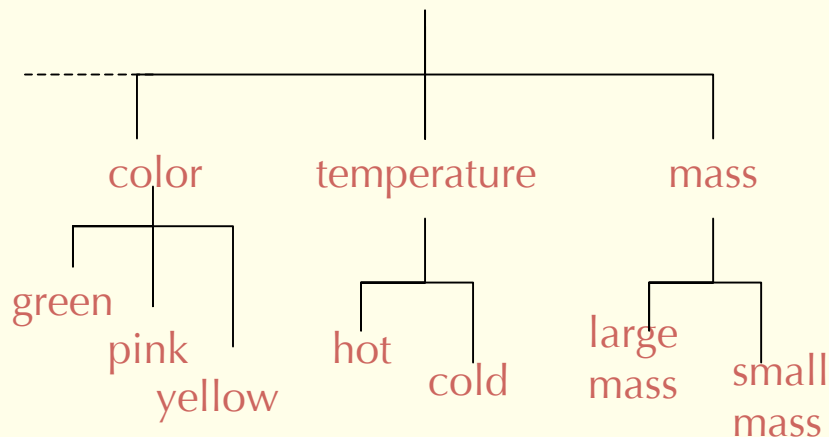
**Note:** *some nodes omitted for brevity*



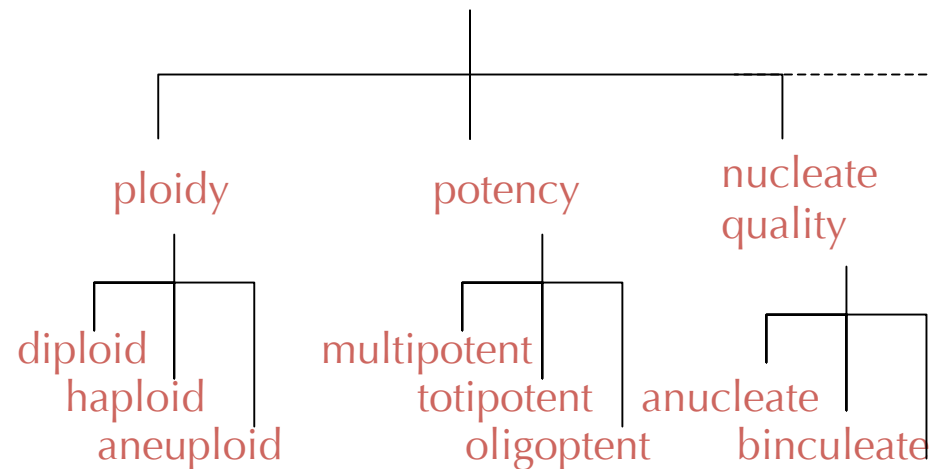
# Divisions by granularity

Monadical quality of a continuant

... Physical quality  
A quality that exists through action of continuants at the physical level of organisation



Cellular quality  
A quality that exists at the cellular level of organisation



# Monadic vs relational

quality of a continuant

