Modern bacterial systematics in practice: Polyphasic taxonomy of the *Burkholderia cepacia* complex

Tom Coenye, PhD
Laboratory for Pharmaceutical Microbiology
Ghent University, Belgium

**Evolution of taxonomy**

- Classification dominated by phenotype

- Classification dominated by genotype
  (“...classification must be a reflection of the natural relationships of bacteria, i.e. their degree of DNA similarity...”)
History of bacterial taxonomy

• 1675: discovery of bacteria by Antonie Van Leeuwenhoek
• 1880s: introduction of “pure culture” techniques (Koch & Petri)
• 1900-1950: taxonomy based on morphology
• 1950s: numerical taxonomy & chemotaxonomy
• 1960s: introduction of genotypical methods
• Present: polyphasic taxonomy

Polyphasic taxonomy

• Integrates several generally accepted ideas for the classification of bacteria
  - Species demarcation is based on DNA-DNA hybridisation experiments
  - Bacterial phylogeny can be deduced from comparative sequence analysis of conserved macromolecules like 16S rDNA
  - It recognises and uses the value of various other methods for distinguishing bacteria at different taxonomic levels

• Goal is to collect as much information as possible in order to generate a useful classification scheme
DNA-DNA hybridisations and the present species concept

Wayne et al., 1987; Ursing et al., 1995

• "... a species generally would include strains with at least 50 – 70% DNA-DNA relatedness ...

• "... a distinct genospecies that cannot be differentiated from another genospecies on the basis of any known phenotypic property should not be named ...

• "... these genospecies could be referred to as genomovars, pending differential biochemical tests ...

16S rDNA sequence analysis

• Present in all bacteria
• Functionally constant
• Mosaic structure of variable and conserved domains
• Comparative 16S rDNA sequence analysis is phylogenetic backbone of modern taxonomy
• Allows to determine phylogenetic neighbourhood of unknown
16S rDNA sequence analysis

- Insufficient diversity for species discrimination (Fox et al., 1992)

- Underestimated diversity in some species
  - 2.5 – 3 % (Stackebrandt & Goebel, 1994)
  - 4.0 – 4.5 % in some \( \varepsilon \)-Proteobacteria (Hanninen et al., 2003)

- Provides a phylogenetic neighbourhood but only a **tentative** identification

Other techniques

- Determine their discriminatory power through the analysis of reference strains previously characterised by 16S sequencing and DDH (“validation”)

- Techniques validated for a particular group can be used for identification in that group:
  - 23S rRNA gene polymorphisms for identification of *Campylobacter* and *Arcobacter* species
  - *recA* gene polymorphisms for identification of *Burkholderia cepacia* complex
  - Identification of *Enterobacteriaceae* by MALDI-TOF MS

- Resolution may be taxon dependent
The genus *Burkholderia*

- Versatile & adapt to changing conditions
- Exploited for:
  - Biocontrol
  - Bioremediation
  - Plant growth promotion
- Safety issues regarding human infections?
- Are the ‘good’ and the ‘bad’ strains the same?
**Burkholderia cepacia**

- Ubiquitous in nature
  - Soil
  - Plants: rhizosphere, roots & shoots
  - Water (including sea water)
  - Humans & animals (infections)
  - Hospital environment
- Conflicting reports about differences between 'environmental' isolates (safe) and 'clinical' isolates (dangerous)
- Strikingly different clinical pictures in CF patients

**Friend: bioremediation**

- Phthalates
- Herbicides (2,4-Dichlorophenoxy acetic acid, 2,4,5-T)
- Trichloroethylene
- Polyaromatic hydrocarbons
- Ether derivatives (gasoline additives)
- …
- Strains: G4, CRE-7, LB400, AC1100, NF100
Foe: onion soft rot

- Universal contaminant (e.g. most frequent bacterial contaminant in cosmetic and pharmaceutical solutions; space shuttle water supply)
- Infections in animals
- Infections in humans
  - nosocomial
  - particular patient groups: cystic fibrosis
Cystic fibrosis

- Hereditary disease caused by a mutation in the CFTR-gene (CF transmembrane conductance regulator)
- Affecting about 1/2500 Caucasians
- Carrier frequency: 1/25
- Malnutrition & chronic lung infection
- Death by decline and failure of lung function
- Traditional pathogens: *S. aureus*, *H. influenzae*, and *P. aeruginosa*

---

*B. cepacia* and CF

- Since 1980ies: rising incidence in *B. cepacia* colonisation
- Strikingly different clinical pictures
  - Stable
  - Slow deterioration
  - 20-35% ‘cepacia syndrome’
- Median life expectancy below 20 y
- Innate multi-resistance
- Highly epidemic strains
- Difficult species-level identification

→ Segregation of Bc+ patients
Median survival by infection (Canada)
(data from M. Corey)

Posttransplant survival %
(Aris et al., 2002, AJRCCM 164:2102-6)
Questions

1. What is the natural biodiversity of these organisms?

2. How can we identify them?

3. Are environmental and clinical *B. cepacia* complex isolates really the same species? The same clone?

Polyphasic taxonomic study of *B. cepacia*

*Why?*

- Conflicting reports about safety of biocontrol strains
- Strikingly different pictures in CF patients
- Differences in epidemiology
- Difficulties associated with diagnosis
Polyphasic taxonomic study of *B. cepacia*

- About 6000 CF, non-CF & environmental isolates
- Reference strains
- Screening methods (whole-cell protein & fatty acid analysis, 16S and *recA* RFLP, AFLP, ribotyping)
- DNA-DNA hybridisation experiments
- 16S rDNA and *recA* sequence analysis
- Biochemical analyses
- International collaborative study - IBCWG
  (http://go.to/cepacia/)

---

Polyphasic taxonomic study of *B. cepacia*

- Screen with same technology to delineate clusters of closely related isolates
- Include representative strains of as many reference species as possible
- Determine phylogenetic position of representatives of unknown clusters
- Compare unknown and closest relatives using DDH in order to find out whether they represent novel species
- Determine optimal identification approach (existing or develop novel)
“B. cepacia” : many species & problems!

- Other bacteria
- B. cepacia → B. cepacia complex

“B. cepacia” : other bacteria

- Many other bacteria are (mis)identified as B. cepacia

- Of 770 isolates referred as B. cepacia, 11% were misidentified - among 281 isolates sent as unidentified or ‘not B. cepacia’, 36% were B. cepacia complex (McMenamin et al. 2000)
“B. cepacia”: other bacteria

- A variety of existing species:
  - *Pseudomonas aeruginosa*
  - *Bordetella hinzii*
  - *Stenotrophomonas maltophilia*
  - *Brevundimonas* sp.
  - *Ralstonia pickettii*
  - *Chryseobacterium* sp.
  - *Achromobacter xylosoxidans*
  - *Enterobacteriacea*
  - *Agrobacterium*
  - *Burkholderia gladioli*
  - ...

“B. cepacia”: other bacteria

- A variety of novel species (mainly β-proteobacteria):
  - *Burkholderia* (B. fungorum, B. caledonica, B. sacchari, B. phymatum, B. tuberum, B. terricola, B. hospita, B. xenovorans,...)
  - *Ralstonia* (R. gilardii, R. mannitolilytica, R. respiraculi, R. taiwanensis)
  - *Pandoraea* (P. sputorum, P. apista, P. pulmonicola, P. pnomenusa, P. norimbergensis, Pandoraea sp.)
  - *Inquilinus limosus*
  - *Herbaspirillum hutiliensis*
  - *Kerstersia gyiorum*
  - *Achromobacter insolitus, Achromobacter spanius*
  - ...

Dr. Tom Coenye, Laboratorium voor Farmaceutische Microbiologie, Universiteit Gent, Workshop Comparative Microbial Genomics and Taxonomy, August 2006
Significance of novel species?

- Clinical significance is mostly unclear
- Several are capable of chronic colonisation
- Transmission
- Erroneous Bc+ status has serious implications for the CF patient
- Erroneous Bc identification result has serious implications for biotechnological applications

B. cepacia → B. cepacia complex

A cluster of multiple genomic species with a high degree of 16S rDNA (98-100%) and recA (94-95%) sequence similarity, and moderate levels of DNA-DNA hybridisation (30-50%), referred to as genomovars.
B. cepacia complex

- `B. cepacia` isolates represent at least 10 genomic species
  - Genomovar I
  - Genomovar II
  - Genomovar III
  - Genomovar IV
  - Genomovar V
  - Genomovar VI
  - Genomovar VII
  - Genomovar VIII
  - Genomovar IX
  - Genomovar X

Biochemical tests (Henry et al. 2001)

- Bcc / Burkholderia / Pandoraea/ Ralstonia

- I-III-VII-VIII-IX / II -VI / IV / V
Whole-cell protein electrophoresis

- Identification at the *B. cepacia* complex level
- Poor discrimination among genomovars I/III/II/VII/VIII/IX
- Best method for rapid identification of ‘other’ species

Whole-cell fatty acid analysis (MIS)

- In general: good ID results at the genus level
- Poor discriminatory power within genera and within the *B. cepacia* complex
- Best alternative if protein method for rapid identification of ‘other’ species fails
16S rDNA

- 98-100% similarity between genomovars

- Enough diversity to develop diagnostic PCR tests for II & V, but not for I, III, IV, VII, VIII, IX (Bauernfeind et al. 1999; LiPuma et al. 1999)

- Ribotyping approach (Brisse et al., 2000)
  - Identifies multiple ribotypes within each genomovar: genomovar identification through ribotype identification
  - Genomovar status of novel ribotypes is always unknown

---

16S rDNA RFLP

RFLP profile Genomovar
(Alu- Cfo- DdeI)

<table>
<thead>
<tr>
<th>AAA</th>
<th>I, III, VII, VIII, IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAB</td>
<td>II</td>
</tr>
<tr>
<td>AAD</td>
<td>II, VII, VIII</td>
</tr>
<tr>
<td>AAC</td>
<td>V</td>
</tr>
<tr>
<td>ABA</td>
<td>IV, VII, VIII, IX</td>
</tr>
<tr>
<td>ABD</td>
<td>IV, IX</td>
</tr>
<tr>
<td>ACC</td>
<td>VI</td>
</tr>
<tr>
<td>BAD</td>
<td>VIII, IX</td>
</tr>
<tr>
<td>BBD</td>
<td>IX</td>
</tr>
</tbody>
</table>
**recA analysis** (Mahenthiralingam et al. 2000)

- 94-95% sequence similarity between genomovars
- Typically 98-99% similarity within genomovars
  - BUT: multiple genomovar I and III alleles
- Sufficient diversity to develop diagnostic PCR tests for each phylogenetic group/allele
- II, IIIA, IIIB, IV, V, VI, VII, VIII: OK
- I: low sensitivity and specificity
- IIIC, IIID, IX: not available

**recA phylogeny**

- B. cepacia
- B. pyrrocinia
- B. cenocepacia IIIC
- B. stabilis
- B. cenocepacia IIIB
- B. anthina
- B. cenocepacia IIIA
- B. ambifaria
- B. cenocepacia IIID
- B. vietnamiensis
- B. dolosa
- B. multivorans
- B. pseudomallei
A phylogeny

**HaeIII-recA RFLP analysis**

- *B. vietnamiensis*
- *B. multivorans*
- Genomovar III
- *B. stabilis*

- Over 70 (!) different RFLP types
Other novel PCR tests

- **bcscQ** and **bcscV**: TTS genes
- **opcL**: outer membrane lipoprotein gene

Questions

1. What is the natural biodiversity of these organisms?
2. How can we identify them? Do we need to?
   - We now understand the phylogenetic structure and relationships of *B. cepacia*-like bacteria and have the tools to investigate the biodiversity
   - Based on this we were able to develop novel identification approaches
3. Are environmental and clinical *B. cepacia* complex isolates really the same species? The same clone?
**B. cepacia complex**

- ‘B. cepacia’ isolates represent at least 9 genomic species
  - Genomovar I: *B. cepacia*
  - Genomovar II: *B. multivorans* sp. nov.
  - Genomovar III: *B. cenocepacia* sp. nov.
  - Genomovar IV: *B. stabilis* sp. nov.
  - Genomovar V: *B. vietnamiensis*
  - Genomovar VI: *B. dolosa* sp. nov.
  - Genomovar VII: *B. ambifaria* sp. nov.
  - Genomovar VIII: *B. anthina* sp. nov.
  - Genomovar IX: *B. pyrrocinia*
  - Genomovar X: *B. ubonensis*

**Questions**

3a. Are “good” and “bad” *B. cepacia* complex strains the same species?

3b. Are “good” and “bad” *B. cepacia* complex strains the same clone?
Prevalence in CF specimens

- Canada (n=449)
- Italy (n=116)
- USA (n=686)
- France (n=171)
- Belgium (n=33)
- Germany (n=174)
- UK (n=17)
- Sweden (n=20)
- Portugal (n=22)

Prevalence in environmental samples

- Balandreau et al. 2001 (AEM 67:982-5)
  - Rhizosphere of wheat and maize, endophytes of wheat and lupine
  - 22 isolates from 2 sites (Australia 19, France 3)
  - III and VII (1 isolate only)

- Fiore et al. 2001 (Env. Microbiol. 3:137-43)
  - Maize rhizosphere
  - 120 isolates from 3 sites
  - I, III, VII, V (1 site - 5 isolates only)
Prevalence in environmental samples

- **Miller et al. 2001** (Abstr. IBCWG meeting)
  - Soil from (sub)urban areas
  - 68 isolates from 3 US cities
  - Mainly VII & IX, and ‘a few’ I and III

- **Gonzalez et al. 2001** (Abstr. IBCWG meeting)
  - Organic soil from onion fields
  - 127 isolates
  - I (31 isolates), III (29), VII (54) and IX (13)

---

Prevalence in environmental samples

- **Pirone et al. 2005** (Env. Microbiol. 7:1734-1742)
  - Maize rhizosphere
  - >600 isolates
  - Mainly IIIB and VII
Overall I, III, VII and IX dominate in soil samples examined. Of these, III dominates in CF patients. Genomovars II, IV, VI and VIII are rarely found in soil samples presently examined. Of these, II dominates in CF patients. Superior capacity of genomovars II and III to colonise CF patients.

Answer 3a (species):

- Knowledge about genomovar distribution in environmental samples is restricted
- Misidentification is common
- Presently 10 B. cepacia complex genomovars, all occurring in human and environmental samples.

The same species that occur in the environment, colonise and infect CF patients, but to varying degrees...
Knowledge about Bcc in various environmental habitats is restricted

**Indirect evidence: segregation works but does not eliminate acquisition!**
- Centers which have Bcc infected patients segregate inpatients and outpatients
- All isolates are typed (RAPD and PFGE)
- Person-to-person transmission is strongly reduced or eliminated, but novel acquisition of strains with unique fingerprints occurs regularly (incidence: 2-6%)

**Direct evidence (1):** Govan et al. ASM News 2000;66:124

- Identical DNA fingerprints (PFGE, RAPD, AFLP) were observed for a recent CF isolate and the genomovar I type strain (onion)!
• PHDC clone: genomovar III epidemic clone colonizing nearly all CF patients in two large treatment centers in the mid-Atlantic region of the USA
• PHDC and several isolates from onion fields (two growth seasons) in the same region had identical DNA fingerprints


• PFGE fingerprinting of environmental and clinical *B. cepacia* complex strains identified multiple identical pairs

Answer 3b (clone):

Human isolates are not necessarily distinct from environmental isolates

Diversity in the genus *Burkholderia*?

- Strongly underestimated!

- Yabuuchi et al. 1992: 7 group II pseudomonads

- *Burkholderia*: extremely versatile genus comprising over 40 species including...
  - Primary human pathogen (eg *B. pseudomallei*)
  - Primary animal pathogen (eg *B. mallei*)
  - Primary plant pathogens (eg *B. glumae*)
  - Soil and plant associated organisms (eg *B. graminis*)
Gut endosymbionts of *Tetraponera* ants

For organisms like *B. cepacia* complex, it is unclear which factors determine ‘virulence’

Correct identification provides a first basis for risk assessment and infection control (‘don’t use *B. cenocepacia* or *B. cepacia complex’…)

Other ‘versatile’ organisms can be found in human clinical samples (eg LB400: *Burkholderia xenovorans* blood isolate)
Conclusions

- *Burkholderia* is a diverse genus with diverse species that live in diverse ecological niches

- The molecular/physiological background of this diversity and adaptability are largely unknown

- Further collaborative molecular taxonomic studies are required to gain further insights into the biodiversity of these genera

- Multiple biotechnologically interesting strains belong to as yet uncharacterised taxa