A Brief Tutorial on Concepts and Terms for Bioinformatic Analysis of Prokaryotic Mobile Genetic Elements (MGEs)

Presented at the Workshop on Sequence Ontology of Mobile Genetic Elements

18-19 Sept 2006
Sanger Center, Hinxton, UK

By

Anne O. Summers, University of Georgia
Athens, GA, USA
Summers@uga.edu

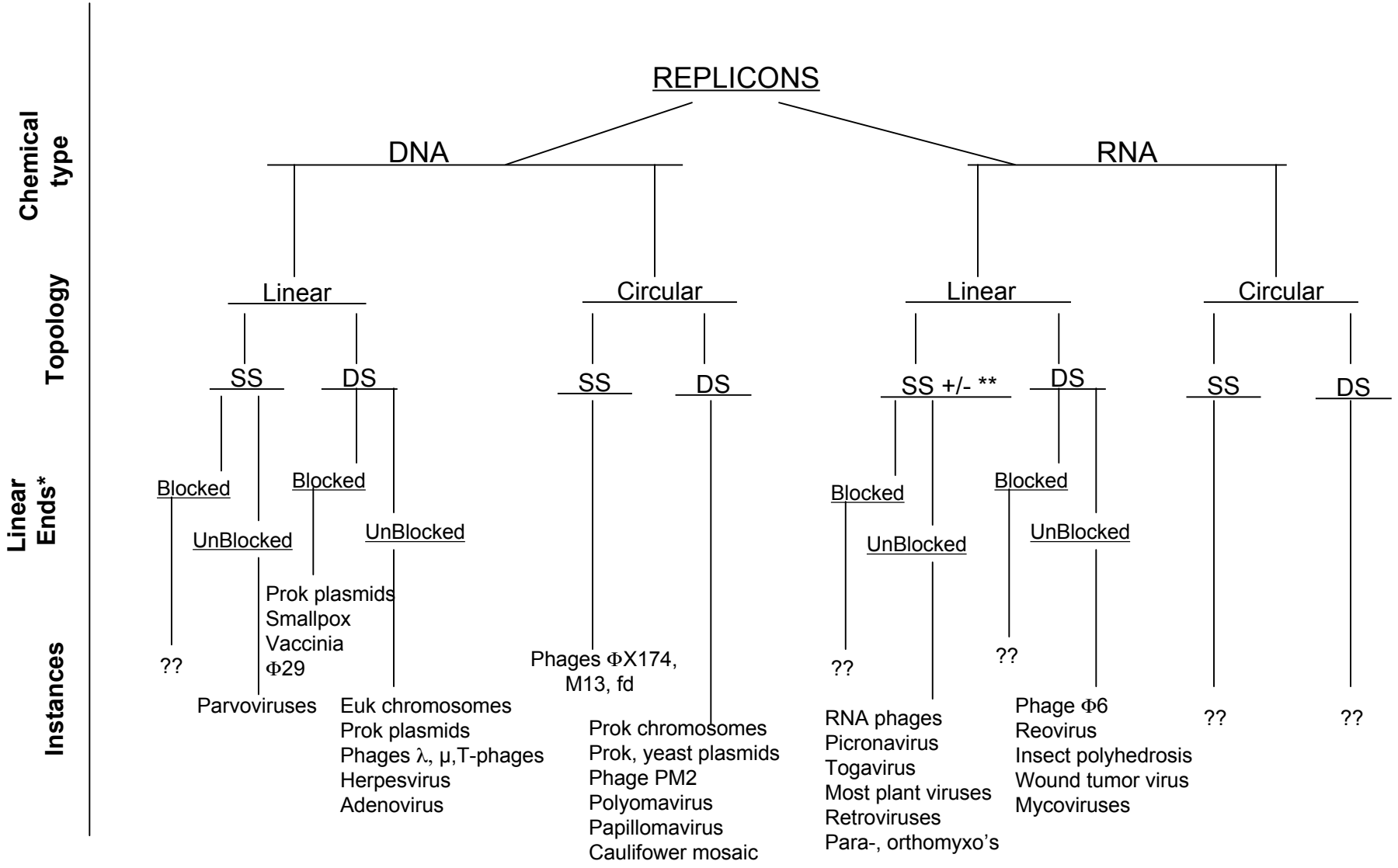
Modes of Gene Transfer in Prokaryotes

- Intercellular
  - Transformation - Transfer of free DNA; does not employ an agent.
  - Transduction - Transfer by adventitious packaging of chromosomal DNA by a phage (an agent).
  - Conjugation - Transfer by a plasmid (an agent) of its own and the chromosome's DNA via cell-to-cell contact.
- Intracellular
  - Transposons and integrons - recombine randomly or site-specifically within a cell but encode no genes for <u>inter</u>cellular transfer.

Plasmids and phages are replicons

- Nucleic acid-based entities which control their own physically independent replication including:
  - “The” main cellular chromosome
  - Plasmids
  - Viruses/bacteriophages
- Pure transposons are not replicons
  - Most transposons only control their own recombination and are replicated by the replicon (plasmid, phage, or chromosome) in which they reside. But see some mosaic forms below.

A HIERARCHY OF REPLICONS BASED ON CHEMISTRY AND TOPOLOGY



*Ends of linear replicons can either be free (unblocked) 5'phosphates and 3' hydroxyls or blocked in short hairpins (in SS) or covalently crosslinked (DS).
** SS RNA replicons can either be directly translatable mRNA's (+) or complementary to mRNA (-), needing to be copied before used as mRNA.

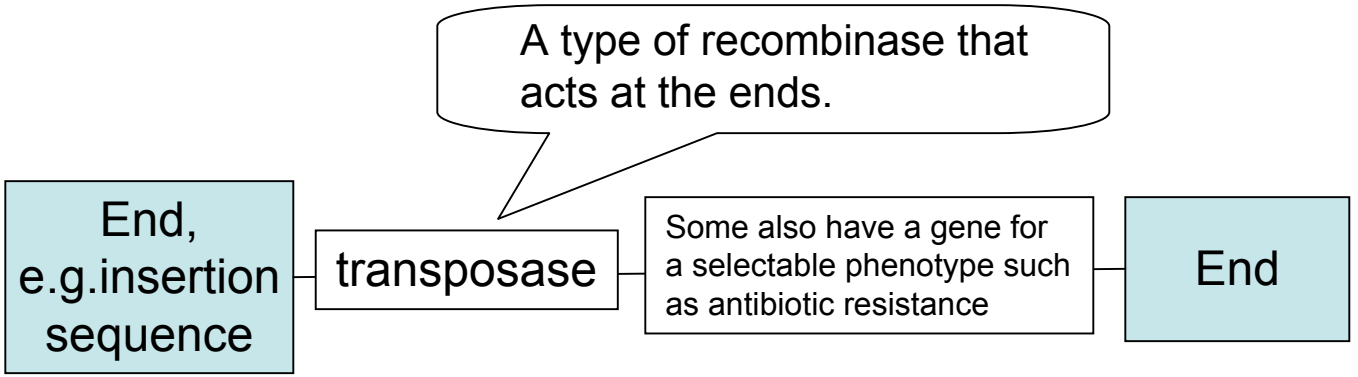
AO Summers, UGA, Nov 2006

Intercellular Gene Transfer by Mobile Replicons

- Plasmids and phages are a subset of DNA replicons that also have genes for self-transfer and for random or site-specific recombination.
- They are not entirely autonomous and variously require some genes of the main chromosome, but their own genes control how often they replicate during the cell division cycle (i.e. copy number).
- Plasmids
  - Currently most studied are circular double-stranded DNA.
  - Transfer is by rolling circle replication, so donor cell keeps copy of plasmid, transfers other copy to recipient cell
  - Size: 1kb - >500 kb; continuum from plasmid < megaplasmid < "mini-chromosome"~1Mb
  - Mid-size: 30-300 kb are typically involved in HGT
- Bacteriophages ('phages)
  - Virulent class – always kill the host cell
  - Temperate class – can kill or reside quiescently in host cell. When it eventually makes more copies of phage and lyses the cell it can accidentally package and transfer host cell DNA.
  - Size: 3-140 kb

Gene Transfer by Transposons

- Can only recombine their own DNA *intra-*cellularly
- Must be on a plasmid or phage to transfer *inter-*cellularly



Basic and Mosaic MGE's

Agents of HGT, MGE (pure prototypes)	Plasmids (RP4, RK2, R100, ColE1)	Phages (lambda)	Transposons (Tn5, Tn10, Tn501)
Mosaic MGE's	P1 A temperate (lysogenic) phage that is a plasmid	Mu A temperate (lysogenic) phage that that replicates by transposition	Conjugative Transposon A transposon that conjugates to other cells, e.g. Tn916
	Genomic Islands have conjugative, phage-derived, and transposon-derived components all located together on the main chromsome. Pathogenicity islands (PAIs) have virulence genes. Also called ICE, integrative conjugative element (20KB - >100s KB)		

MGE's have 3 defining processes

		1. Self-Replication	2. Transfer	3. Site-specific Recombination
Genes	Trans-acting: Enzymes and NA binding proteins	Replicases, NA polymerases, rolling circle replication, phage 'host-takeover' functions	DNA "pump", a Type IV secretion system (T4SS). Phage packaging proteins	Transposases, integrases, excisases, resolvases
	Cis-acting: NA interaction or cleavage sites	oriV	oriT, phage packaging sites, cos site	Gene cassettes, Att sites, Tnp IR's, Res sites

MGE's also carry "baggage" loci: functions valuable to the host cells but not intrinsic to being a gene transfer agent.

Examples of "baggage" genes		**1. Pathogenicity**	**2. Metabolism**	**3. Resistance**
Genes	Trans-acting: Enzymes, transport proteins and NA binding proteins	Toxins , invasion proteins, colonization factors	Catabolic pathways for xenobiotic compounds. Nitrogen fixation, photosynthesis	Antibiotic resistance, toxic metal resistance.
	Cis-acting: NA interaction sites	Relevant operators, promoters, etc.	Relevant operators, promoters, etc	Relevant promoters, operators, etc

Some "baggage" loci reside within transposons that are carried by plasmids or phages.

Ontology Needs & Benefits for Mobile Genetic Elements

- **Missing gene ontologies for intrinsic MGE functions, e.g.:**
  - DNA pump (variant of general secretion systems)
  - Recombinases (transposases, resolvases, integrases)
  - Phage structural/packaging genes
    - Tail fibers
    - Capsid proteins (self-assemble into phage protein coat)
- **Missing sequence ontologies for intrinsic MGE functions, e.g.:**
  - Transfer origins, *oriT* and unique replication origins, *oriV*
  - *Res* sites for transposition
  - *Att* sites for integration
  - Phage packaging sites
- **Benefits of improved MGE ontology for the community:**
  - Literature searching
  - Find novel relationships, insights
  - Facilitate expert annotation - A.I. makes it easier and faster to incorporate their expertise