DISCOVERY OF NOVEL DRUG TARGETS AGAINST PATHOGENIC PROTOZOA: THE PROMISE OF METABOLIC RECONSTRUCTION

BIOINFORMATICS FOR AFRICA 2007 CONFERENCE AND TRAINING
30TH MAY 2007

PRESENTED BY: LILLIAN WAMBUA
ILRI, Kenya.

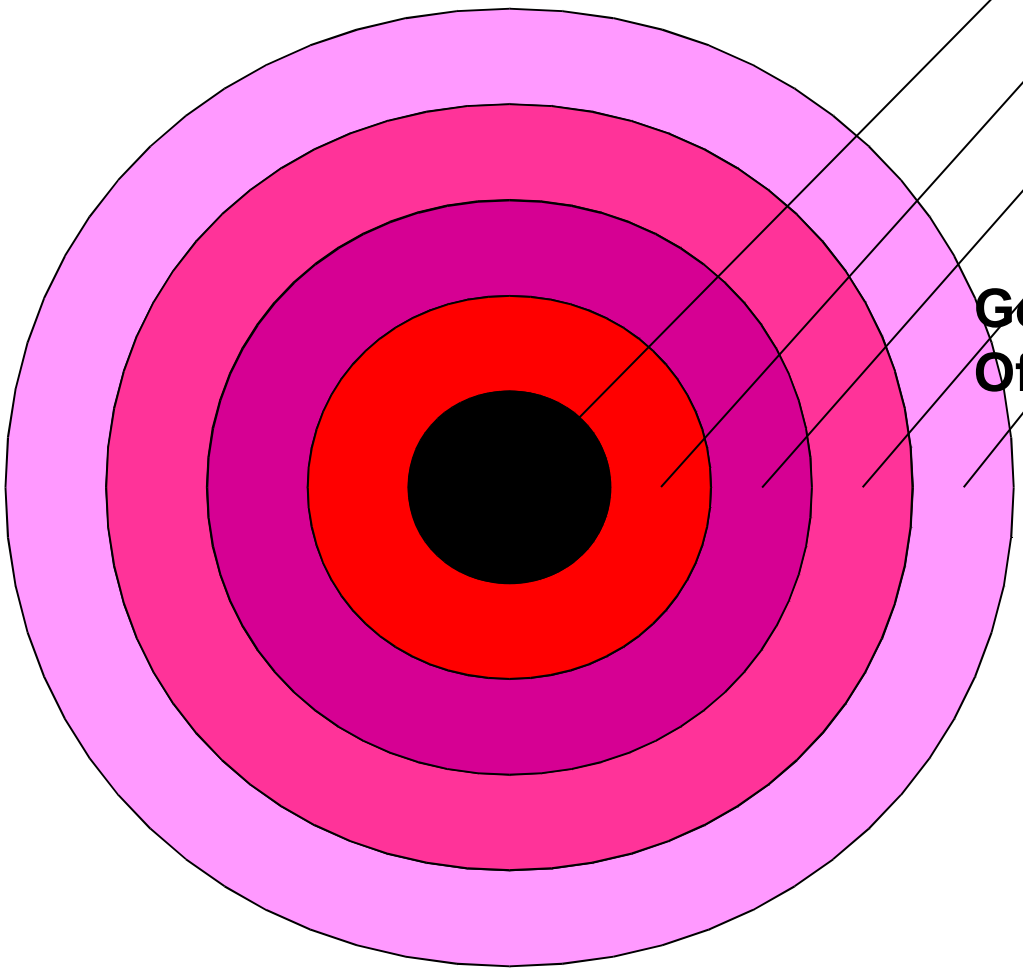
UNIVERSITY OF LEEDS

ILRI
INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE

CURRENT STATUS

- Protozoan diseases
- Drug resistance and toxicity
- Dire need for new drugs
- Need to discover novel drug targets

Research approaches in drug discovery



Metabolism is a rich target resource

- DNA – genes – proteins – networks - systems
- Metabolism is key to any organism
- Perturbation of metabolism affects survival
- Metabolic enzymes are attractive targets
- Present in pathogen only; substantially divergent from host; degree of differential dependence on enzyme/pathway

Problematic Protozoans

- Considered most primitive eukaryotes
- Acquired specialized organelles/pathways – apicoplast, mitochondria and peroxisomes
- Pathogenic protozoans have uniquely undergone subsequent loss or gain to survive or adapt
- Represent highly divergent organisms – present problems in annotation – pathway holes

Introducing metaSHARK

- **Meta**bolic **S**earc**H** **A**nd **R**econstruction **K**it
- Automatic software for prediction of metabolic enzymes and networks
- Developed by Leeds University

  http://www.bioinformatics.leeds.ac.uk/shark/
- Detects enzymes in un-annotated data

Why metaSHARK

- Does not require annotated data i.e. protein predictions
- Input is raw sequences: Genomic, ESTs or GSS
- Useful for annotation and extraction of metabolism information from on-going or complete genomes
- Detects close as well as distant enzyme homologs

MetaSHARK gene-finding

metaSHARK Incorporates 2 components:

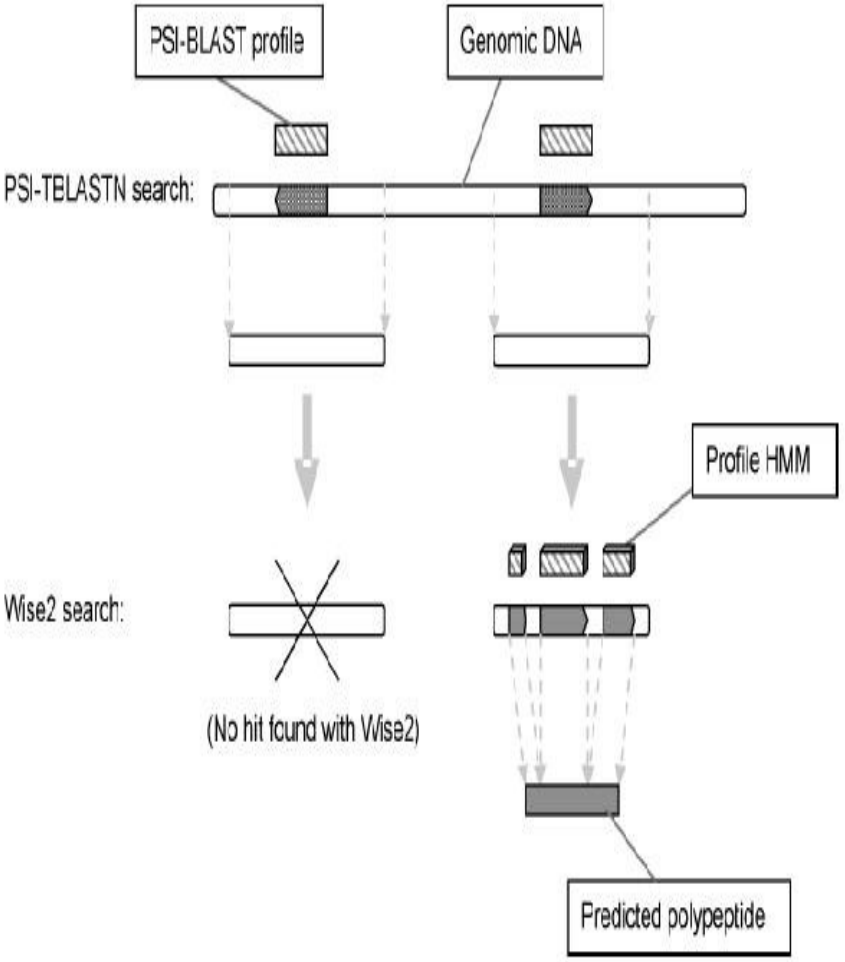
- Gene-detection package i.e. SHARKhunt
- Visualization package i.e. SHARKview

Gene-finding SHARKhunt utilizes:

- PRIAM profiles
- Associated HMM profiles
- DNA sequences – genomic, ESTs or GSSs

Gene-finding with metaSHARK

- PSI-TBLASTN against PRIAM (filter step)
- Wise2 alignment to HMMs
- Predicted polypeptide
- Realigned to original PRIAM profile
- Final E-value score



Scope of study

Metabolic Pathways studied:

- Energy metabolism
  - Glycolysis
  - TCA cycle
- Amino acid biosynthesis
  - Chorismate synthesis via the shikimate pathway
- Biosynthesis of co-factors
  - Synthesis Pantothenate and CoA
  - Folate synthesis

Scope of study

Pathogenic protozoa under investigation

Apicomplexa

- Plasmodium - Malaria
- Toxoplasma - Toxoplasmosis
- Eimeria – Avian Coccidiosis
- Cryptosporidium - Cryptosporidiosis
- Theileria – East coast fever
- Babesia – Babesiosis/Red water disease

Kinetoplastida

- Leishmania - Leishmaniasis
- Trypanosoma – Sleeping sickness, Nagana and Chaga's disease

Amoebidae

- Entamoeba – Amoebic dysentry

Approach

1. Database searches on KEGG, Metacyc, PUMA2

2. Metabolic reconstruction with metaSHARK
   - Genome downloads
   - Gene detection by SHARKhunt
   - Enzymes assertions
   - E-Value 1 X 10^{-10}

3. Manual rebuilding of pathways of interest

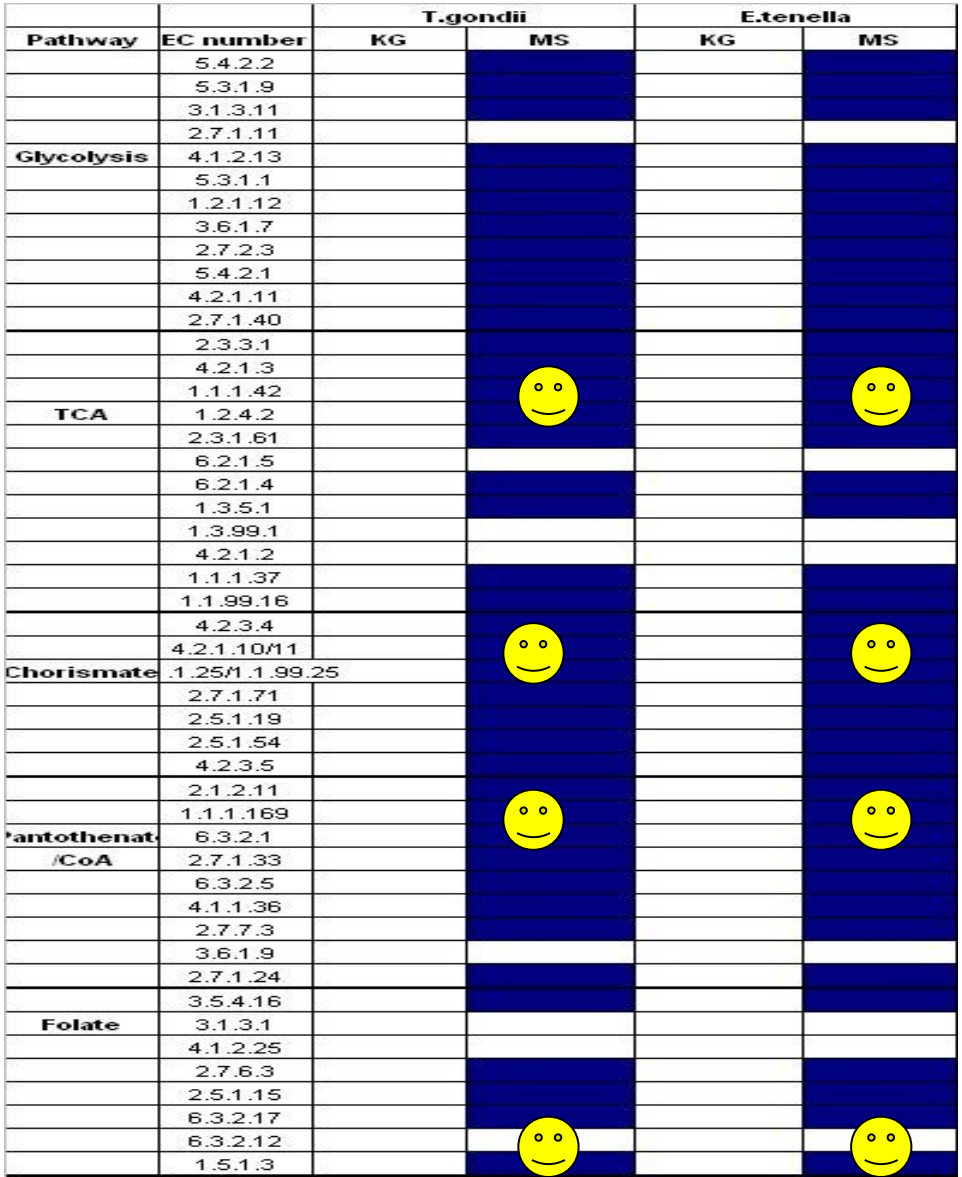
4. Comparative analysis

Findings

C.Parvum and E.histolytica

		KG	MS	KG	MS
Pathway	EC number	C.parvum		E.histolytica	
	5.4.2.2	■	■	■	■
	5.3.1.9	■	■	■	■
	3.1.3.11				
	2.7.1.11	■		■	■
Glycolysis	4.1.2.13	■	■	■	■
	5.3.1.1		■		■
	1.2.1.12	■	■	■	■
	3.6.1.7		■		
	2.7.2.3	■	■	■	■
	5.4.2.1	■	■	■	■
	4.2.1.11	■	■	■	■
	2.7.1.40	■	■	■	■
	2.3.3.1				
	4.2.1.3				
	1.1.1.42				
TCA	1.2.4.2		☹		☹
	2.3.1.61				
	6.2.1.5				
	6.2.1.4				■
	1.3.5.1				
	1.3.99.1				
	4.2.1.2			■	
	1.1.1.37		■	■	■
	1.1.99.16				
	4.2.3.4				
	4.2.1.10/11	■	■	■	■
Chorismate	1.1.1.25/1.1.99.25				
	2.7.1.71				
	2.5.1.19				
	2.5.1.54				
	4.2.3.5				
	2.1.2.11				
	1.1.1.169				
Pantothenate	6.3.2.1				
/CoA	2.7.1.33				■
	6.3.2.5		■		
	4.1.1.36		■	■	■
	2.7.7.3			■	
	2.7.1.24		■	■	■
Folate	3.1.3.1				
	4.1.2.25				
	2.7.6.3				
	2.5.1.15				
	6.3.2.17				
	6.3.2.12				
	1.5.1.3	■	■		

E.Tenella and T.gondii

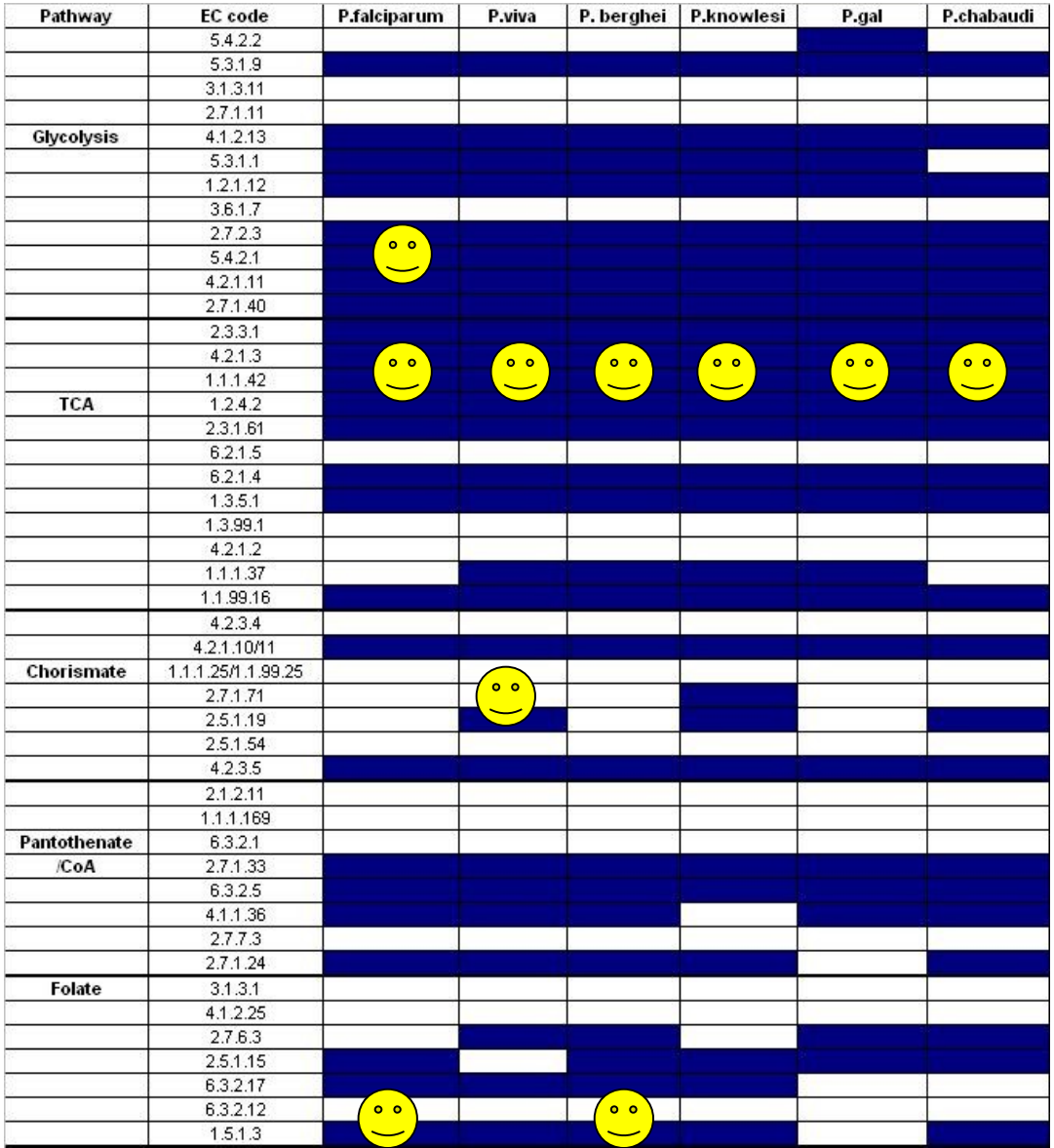


		T.gondii		E.tenella	
Pathway	EC number	KG	MS	KG	MS
	5.4.2.2				
	5.3.1.9				
	3.1.3.11				
	2.7.1.11				
Glycolysis	4.1.2.13				
	5.3.1.1				
	1.2.1.12				
	3.6.1.7				
	2.7.2.3				
	5.4.2.1				
	4.2.1.11				
	2.7.1.40				
	2.3.3.1				
	4.2.1.3				
	1.1.1.42				
TCA	1.2.4.2				
	2.3.1.61				
	6.2.1.5				
	6.2.1.4				
	1.3.5.1				
	1.3.99.1				
	4.2.1.2				
	1.1.1.37				
	1.1.99.16				
	4.2.3.4				
	4.2.1.10/11				
Chorismate	.1.25/1.1.99.25				
	2.7.1.71				
	2.5.1.19				
	2.5.1.54				
	4.2.3.5				
	2.1.2.11				
	1.1.1.169				
antothenat	6.3.2.1				
/CoA	2.7.1.33				
	6.3.2.5				
	4.1.1.36				
	2.7.7.3				
	3.6.1.9				
	2.7.1.24				
	3.5.4.16				
Folate	3.1.3.1				
	4.1.2.25				
	2.7.6.3				
	2.5.1.15				
	6.3.2.17				
	6.3.2.12				
	1.5.1.3				

The tritryps

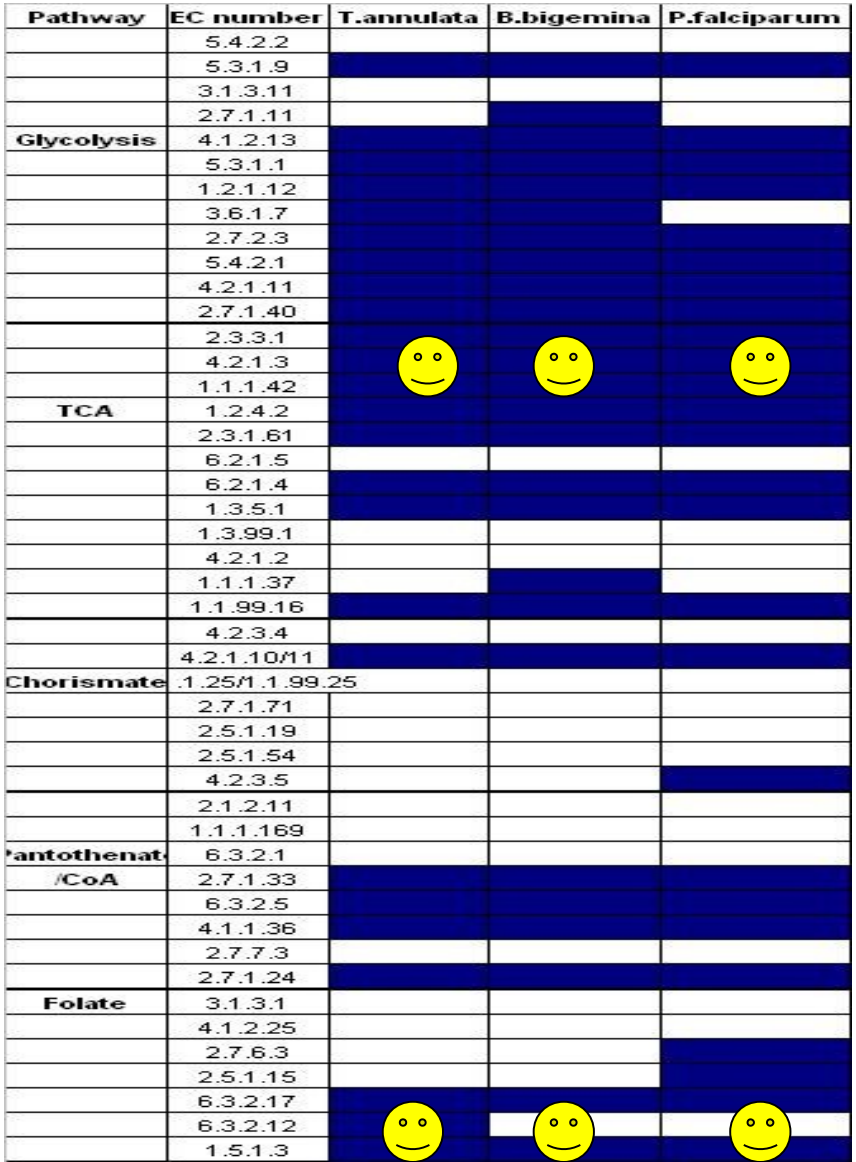
		Leishmania		T.brucei		T.cruzi	
Pathway	EC code	KEGG	MetaSHARK	KEGG	MetaSHARK	KEGG	MetaSHARK
	5.4.2.2						
	5.3.1.9						
	3.1.3.11						
	2.7.1.11						
Glycolysis	4.1.2.13						
	5.3.1.1						
	1.2.1.12						
	3.6.1.7						
	2.7.2.3						
	5.4.2.1						
	4.2.1.11						
	2.7.1.40						
	2.3.3.1						
	4.2.1.3						
	1.1.1.42						
TCA	1.2.4.2						
	2.3.1.61						
	6.2.1.5						
	6.2.1.4						
	1.3.5.1						
	1.3.99.1						
	4.2.1.2						
	1.1.1.37						
	1.1.99.16						
	4.2.3.4						
	4.2.1.10/11						
Chorismate	1.1.1.25/1.1.99.25						
	2.7.1.71						
	2.5.1.19						
	2.5.1.54						
	4.2.3.5						
	2.1.2.11						
	1.1.1.169						
Pantothenate	6.3.2.1						
/CoA	2.7.1.33						
	6.3.2.5						
	4.1.1.36						
	2.7.7.3						
	2.7.1.24						
Folate	3.1.3.1						
	4.1.2.25						
	2.7.6.3						
	2.5.1.15						
	6.3.2.17						
	6.3.2.12						
	1.5.1.3						

The Plasmodia



Pathway	EC code	P.falciparum	P.viva	P. berghei	P.knowlesi	P.gal	P.chabaudi
	5.4.2.2						
	5.3.1.9						
	3.1.3.11						
	2.7.1.11						
Glycolysis	4.1.2.13						
	5.3.1.1						
	1.2.1.12						
	3.6.1.7						
	2.7.2.3						
	5.4.2.1						
	4.2.1.11						
	2.7.1.40						
	2.3.3.1						
	4.2.1.3						
	1.1.1.42						
TCA	1.2.4.2						
	2.3.1.61						
	6.2.1.5						
	6.2.1.4						
	1.3.5.1						
	1.3.99.1						
	4.2.1.2						
	1.1.1.37						
	1.1.99.16						
	4.2.3.4						
	4.2.1.10/11						
Chorismate	1.1.1.25/1.1.99.25						
	2.7.1.71						
	2.5.1.19						
	2.5.1.54						
	4.2.3.5						
	2.1.2.11						
	1.1.1.169						
Pantothenate	6.3.2.1						
/CoA	2.7.1.33						
	6.3.2.5						
	4.1.1.36						
	2.7.7.3						
	2.7.1.24						
Folate	3.1.3.1						
	4.1.2.25						
	2.7.6.3						
	2.5.1.15						
	6.3.2.17						
	6.3.2.12						
	1.5.1.3						

Babesia and Theileria



Pathway	EC number	T.annulata	B.bigemina	P.falciparum
	5.4.2.2			
	5.3.1.9			
	3.1.3.11			
	2.7.1.11			
Glycolysis	4.1.2.13			
	5.3.1.1			
	1.2.1.12			
	3.6.1.7			
	2.7.2.3			
	5.4.2.1			
	4.2.1.11			
	2.7.1.40			
	2.3.3.1			
	4.2.1.3			
	1.1.1.42			
TCA	1.2.4.2			
	2.3.1.61			
	6.2.1.5			
	6.2.1.4			
	1.3.5.1			
	1.3.99.1			
	4.2.1.2			
	1.1.1.37			
	1.1.99.16			
	4.2.3.4			
	4.2.1.10/11			
Chorismate	.1.25/1.1.99.25			
	2.7.1.71			
	2.5.1.19			
	2.5.1.54			
	4.2.3.5			
	2.1.2.11			
	1.1.1.169			
antothenat	6.3.2.1			
/CoA	2.7.1.33			
	6.3.2.5			
	4.1.1.36			
	2.7.7.3			
	2.7.1.24			
Folate	3.1.3.1			
	4.1.2.25			
	2.7.6.3			
	2.5.1.15			
	6.3.2.17			
	6.3.2.12			
	1.5.1.3			

Take home message

- Need for more powerful annotation software for protozoa
- There are numerous targets awaiting discovery and experimental validation.
- Study of metabolism is required to elucidate new drug targets, promote understanding of biology of parasites which will give great insights to host pathogen interactions.
- Tight link btwn evolution and metabolism.

ACKNOWLEDGEMENTS

ILRI
INTERNATIONAL LIVESTOCK RESEARCH INSTITUTE

- Mentors
- Friends and Family

UNIVERSITY OF LEEDS

- Prof Dave Westhead
- Dr. Glenn McConkey
- Bioinformatics team in Leeds

- The Marie Curie foundation for an Early Stage Training Fellowship

Thank you!!