

Harnessing high throughput technologies to tackle neglected parasitic diseases of major socioeconomic importance



THE UNIVERSITY OF  
MELBOURNE

Parasite Genetics and Genomics  
Faculty of Veterinary Science  
The University of Melbourne

# Parasitic helminths



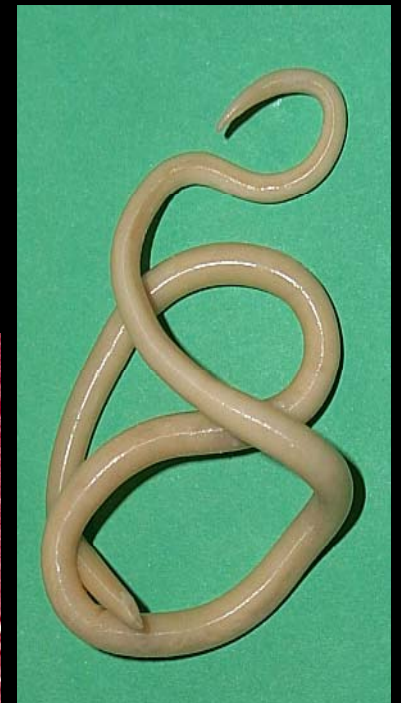
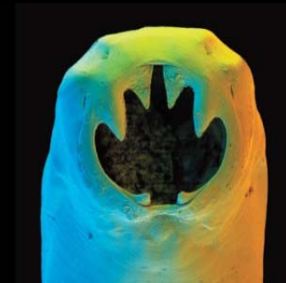
5 mm





# Major advances in the genomics of soil transmitted nematodes

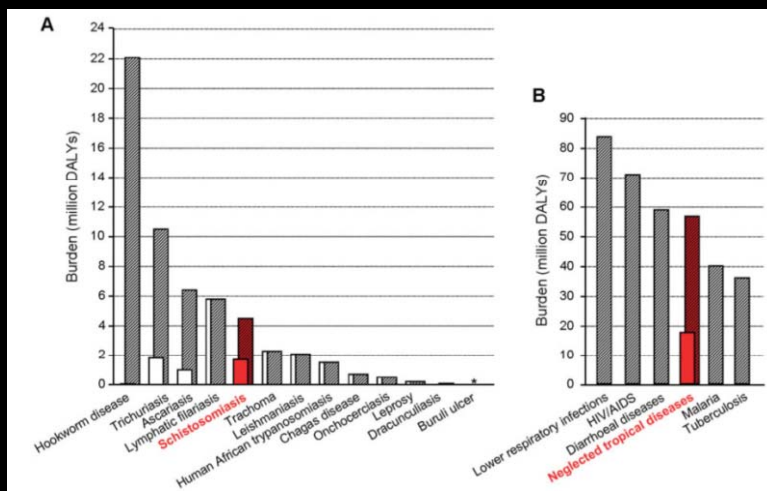
Aaron Jex  
Faculty of Veterinary Science  
ajex@unimelb.edu.au



# 'The great plague of mankind'

## The world's neglected helminths

- Schistosomes – 250 million
- Food-borne trematodes – 80 million
- Taeniid cestodes – >50 million
- **Soil transmitted helminths – ~2 billion**
- Filarial nematodes – 160 million



Developing vaccines to combat hookworm infection and intestinal schistosomiasis

Peter J. Hotez\*, Jeffrey M. Beatty and Alex Loukas<sup>§</sup>

Neglected Diseases

**Hookworm: "The Great Infection of Mankind"**

Peter J. Hotez\*, Jeff Bethony, Maria Elena Bottazzi, Simon Brooker, Paulo Buss

Open access, freely available online

Rescuing the bottom billion through control of neglected tropical diseases

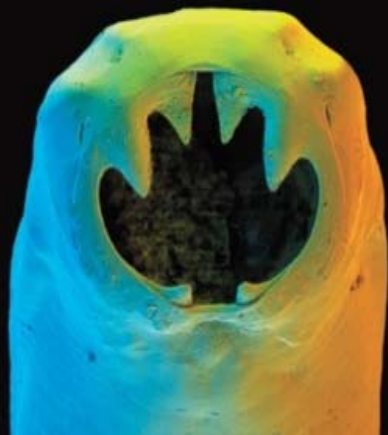
Peter J Hotez, Alan Fenwick, Loren

**The *Schistosoma japonicum* genome reveals features of host-parasite interplay**

The *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium\*

# Soil transmitted helminthiases

- *Ascaris lumbricoides* – 1.2 billion
- *Necator americanus*
- *Ancylostoma duodenale* } 750 million
- *Trichuris trichiura* – > 800 million
- *Strongyloides stercoralis* – 100 million??



# The draft genome for *Ascaris suum*: a major new resource of a key neglected parasite

## LETTER

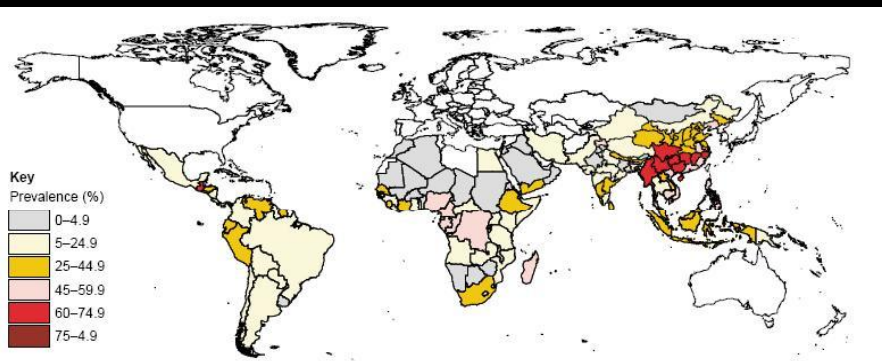
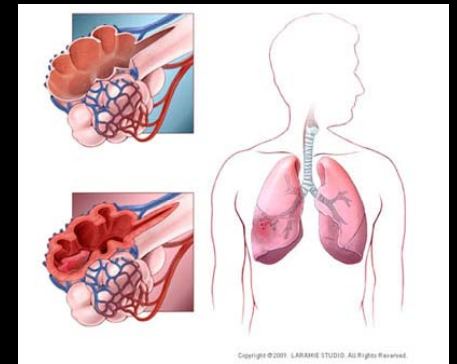
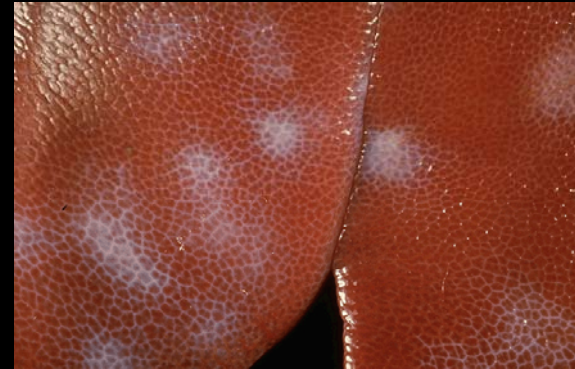
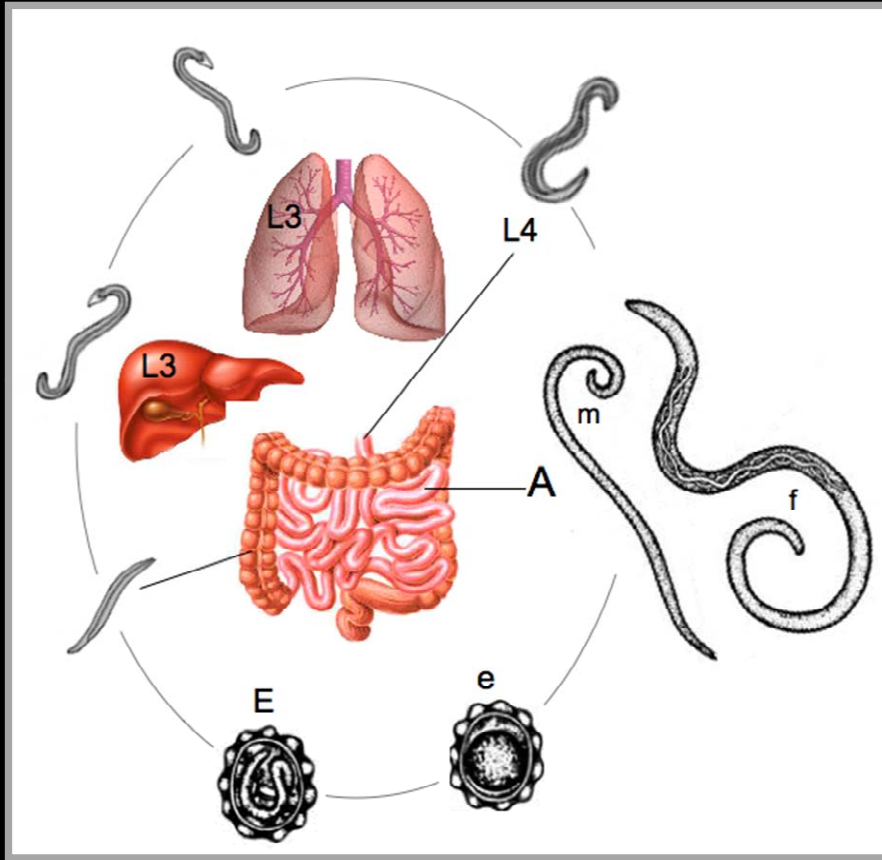
doi:10.1038/nature10553

### *Ascaris suum* draft genome

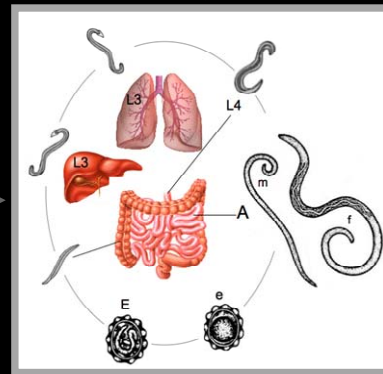
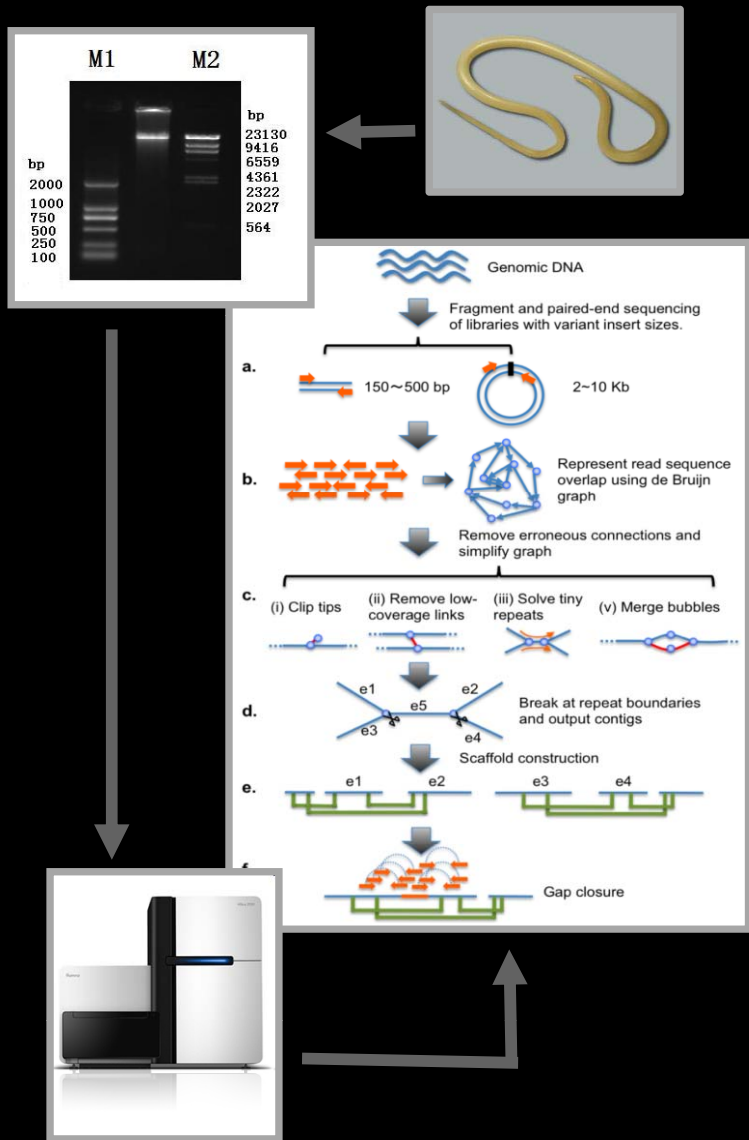
Aaron R. Jex<sup>1\*</sup>, Shiping Liu<sup>2\*</sup>, Bo Li<sup>2\*</sup>, Neil D. Young<sup>1\*</sup>, Ross S. Hall<sup>1</sup>, Yingrui Li<sup>2</sup>, Linfeng Yang<sup>2</sup>, Na Zeng<sup>2</sup>, Xun Xu<sup>2</sup>, Zijun Xiong<sup>2</sup>, Fangyuan Chen<sup>2</sup>, Xuan Wu<sup>2</sup>, Guojie Zhang<sup>2</sup>, Xiaodong Fang<sup>2</sup>, Yi Kang<sup>2</sup>, Garry A. Anderson<sup>1</sup>, Todd W. Harris<sup>3</sup>, Bronwyn E. Campbell<sup>1</sup>, Johnny Vlaminc<sup>4</sup>, Tao Wang<sup>4</sup>, Cinzia Cantacessi<sup>1</sup>, Erich M. Schwarz<sup>5</sup>, Shoba Ranganathan<sup>6</sup>, Peter Geldhof<sup>4</sup>, Peter Nejsum<sup>7</sup>, Paul W. Sternberg<sup>5</sup>, Huanming Yang<sup>2</sup>, Jun Wang<sup>2</sup>, Jian Wang<sup>2</sup> & Robin B. Gasser<sup>1</sup>



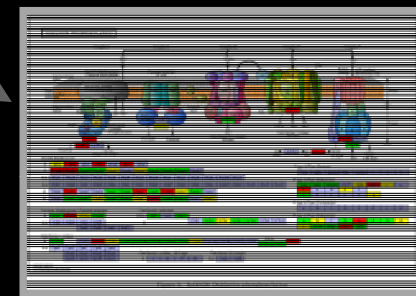
# Ascaris and ascariasis



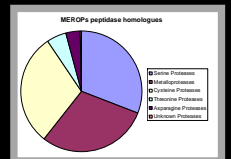
# Sequencing the *A. suum* genome



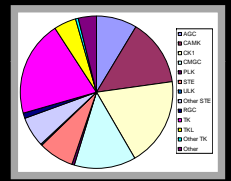
KEGG/KOBAS



MEROPs



KS-Sarfari





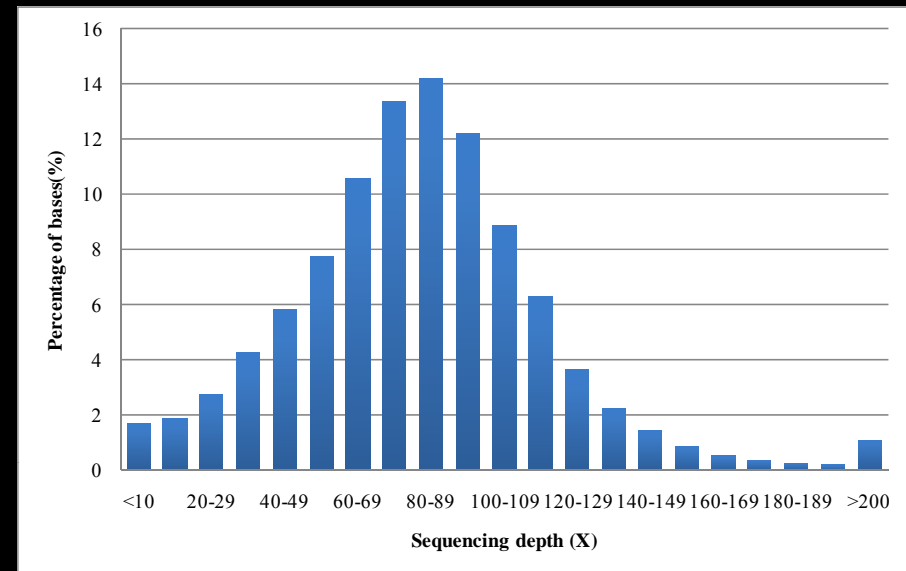
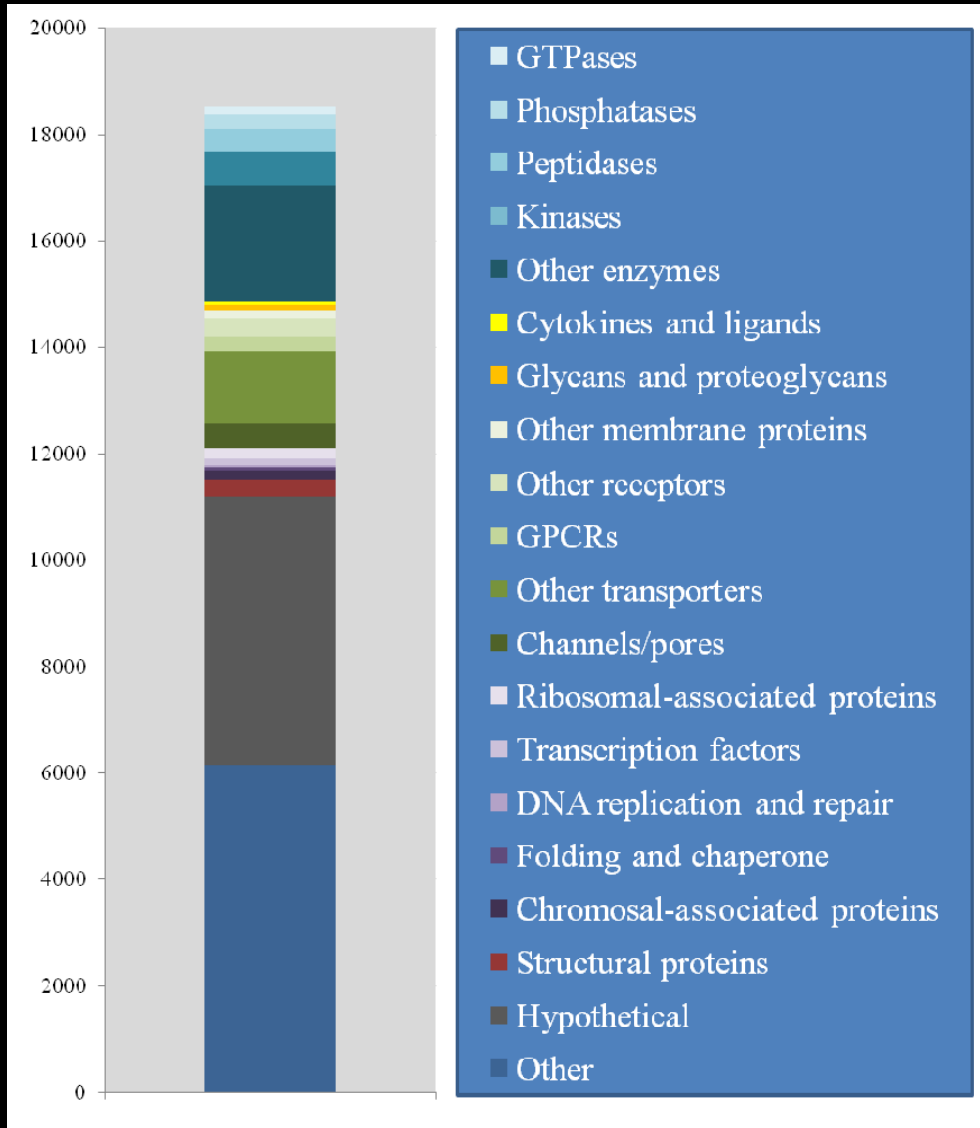
# Genome summary

Genome size: 273 Mb (N50: 408 kb)

GC content: 38%

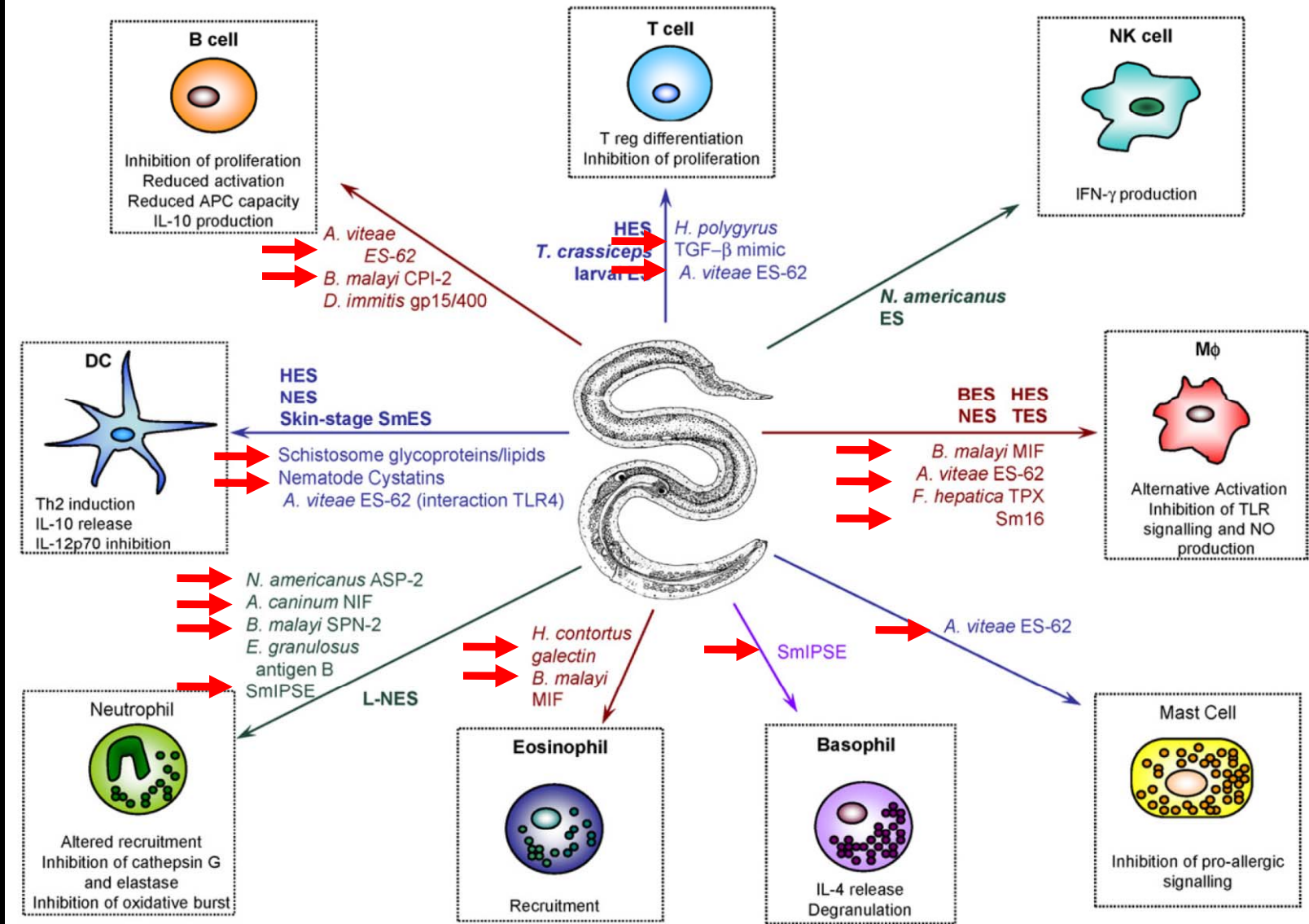
Repetitive sequence: 4%

Genes: 18,542



# Immunomodulation

J.P. Hewitson et al. / Molecular & Biochemical Parasitology 167 (2009) 1–11



# Differential transcription

## Migrating larvae

Secreted peptidases

e.g., C1, C2A, S9, S33, M1 and M12

Olfactory chemosensory molecules

e.g., *egl-4*, *grk-2*, *kin-29*, *tax-4*, *odr-1*, 3, 4 and 10

Immunomodulators

e.g., c-type lectins, serpins, cystatins, oxidoreductases, *pafa* and *mif-4*

## Female reproductive tissue

Egg laying / oogenesis

e.g., *cat-1*, *chs-2*, *unc-54*, *cbd-1* and *pqn-74*

Vulva development

e.g., *noah-1*, *nhr-25*, *cog-1* and *pax-3*

Embryogenesis

e.g., *cam-1* and *unc-6*

## Male reproductive tissue

Sperm / spermatogenesis

e.g., *fer-1*, *spe-4*, 6, 9, 10, 15 and 41, *alg-4* and *msp-2*, 3, 57, 78

Kinases

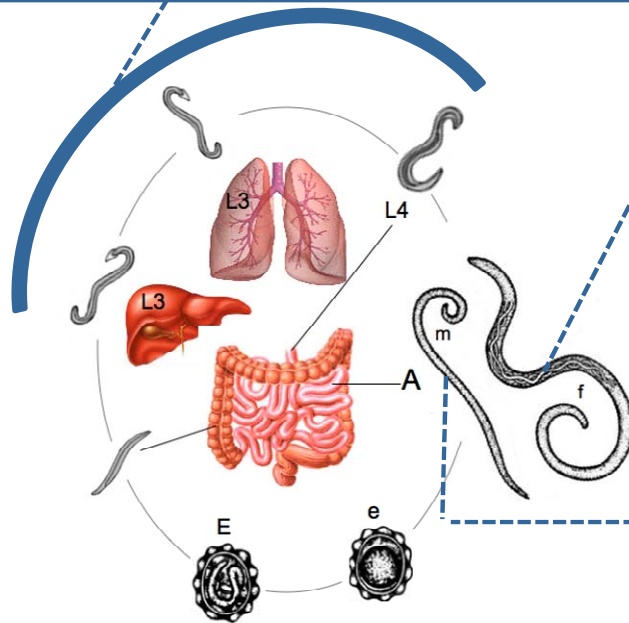
e.g., CK1s and CAMKs

Phosphatases

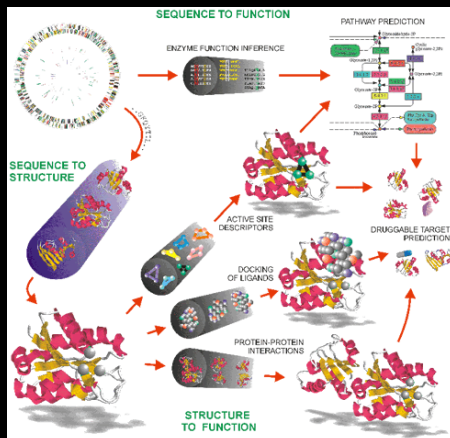
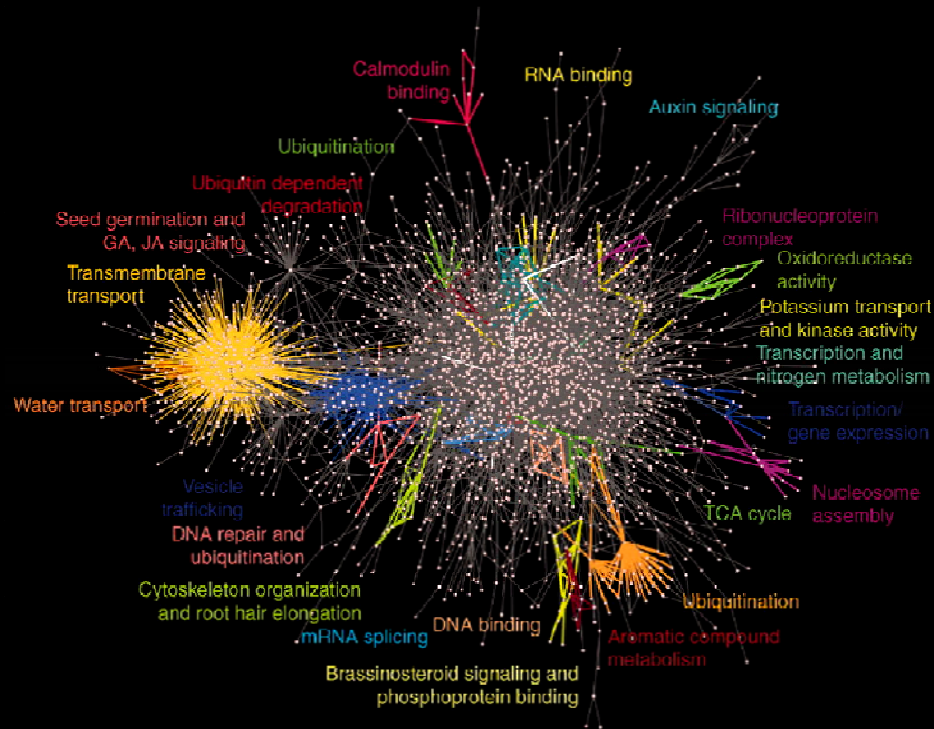
e.g., PTPs and STPs

Peptidases

predominantly metalloproteases



# Druggable genome



Predicted targets: 204

Single copy 'essential' chokepoints

- CDP-diacylglycerol-inositol 3-phosphatidyltransferase
- G protein-coupled receptor kinase 5
- Phosphoribosylformylglycinamide synthase
- **Inosine-5'-monophosphate dehydrogenase**
- Phospho-N-acetylmuramoyl-pentapeptide-transferase

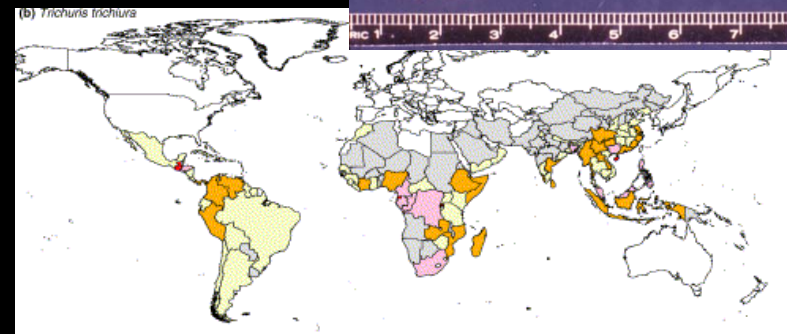
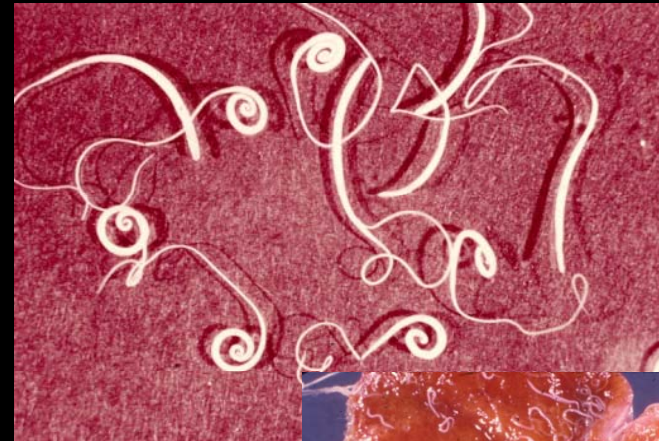
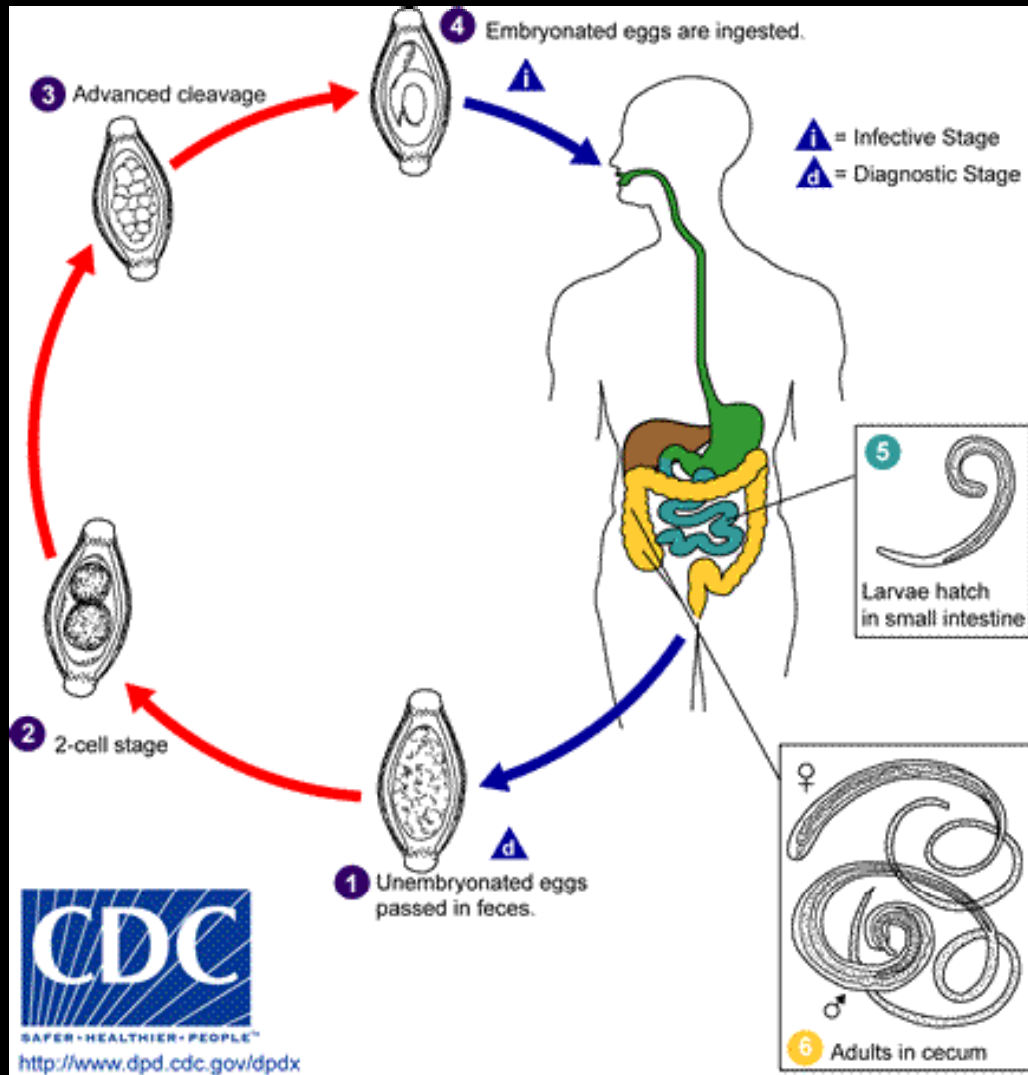
# The *Trichuris suis* draft genome

Aaron R. Jex<sup>1\*</sup>, Peter Nejsum<sup>2</sup>, Erich Schwarz<sup>1,3</sup>, Li Hu<sup>4</sup>, Neil D. Young<sup>1</sup>, Ross S. Hall<sup>1</sup>, Pasi K. Korhonen<sup>1</sup>, Shengguang Liao<sup>4</sup>, Stig Thamsborg<sup>2</sup>, Jinquan Xia<sup>4</sup>, Pengwei Xu<sup>4</sup>, Shaowei Wang<sup>4</sup>, Jean-Pierre Scheerlink<sup>1</sup>, Andreas Hoffman<sup>5</sup>, Paul W. Sternberg<sup>6</sup>, Jun Wang<sup>4</sup> & Robin B. Gasser<sup>1</sup>

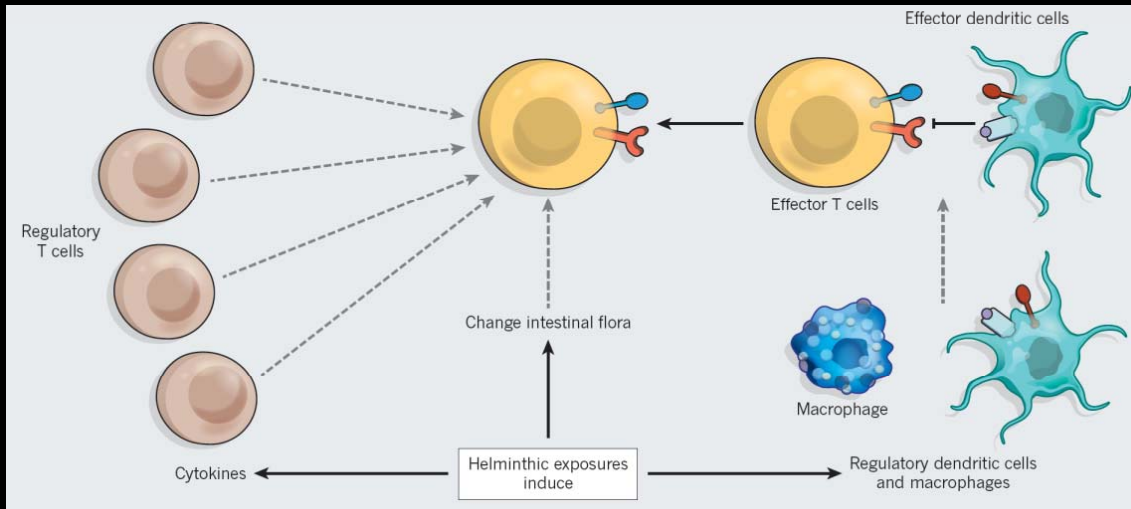
1. The University of Melbourne, Australia
2. The University of Copenhagen, Denmark
3. Cornell University, USA
4. Beijing Genomics Institute, China
5. Griffith University, Australia
6. California Institute of Technology, USA



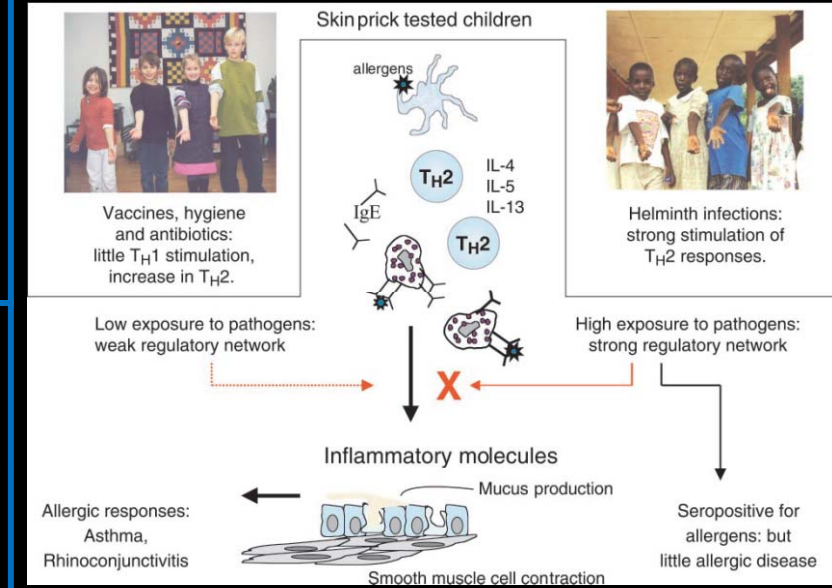
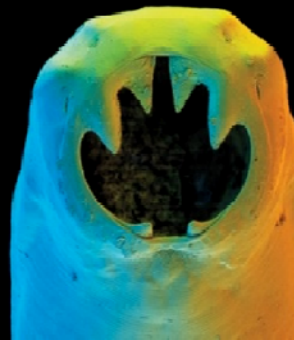
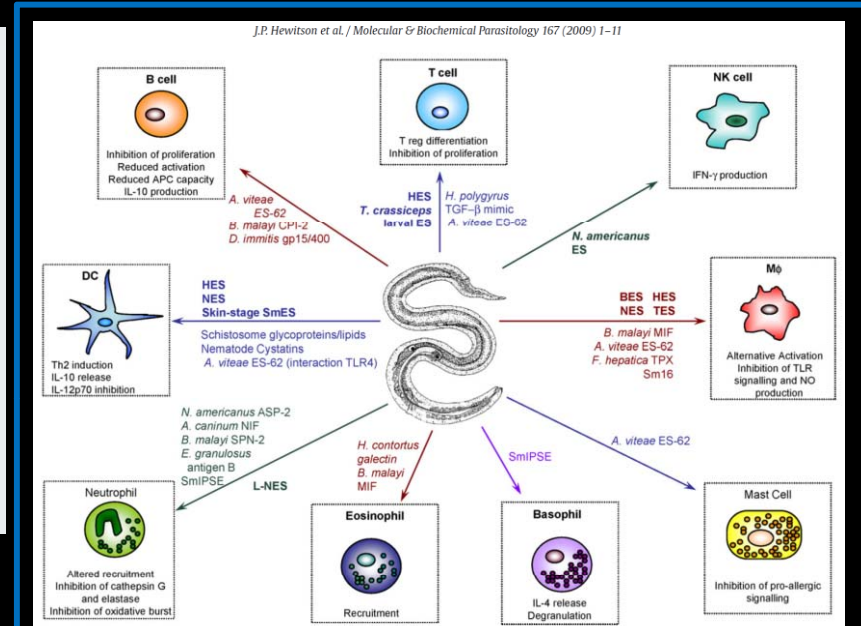
# Trichuriasis - disease



# Helminth exposure vs immunoregulation



Weinstock et al, Nature, 2012



# T. suis ova

Summers et al, Gastroenterol, 2005

## TSO (*Trichuris suis* ova or CNDO-201)

Indication	Pre-Clinical	Phase 1	Phase 2a*	Phase 2	Phase 3
Crohn's Disease	→				
Ulcerative Colitis	→				
Multiple Sclerosis	→				
Autism	→				
Psoriasis	→				
Type-1 Diabetes	→	2H2013 →	→		
Psoriatic Arthritis	→	2H2013 →	→		
Rheumatoid Arthritis	→	2H2013 →	→		

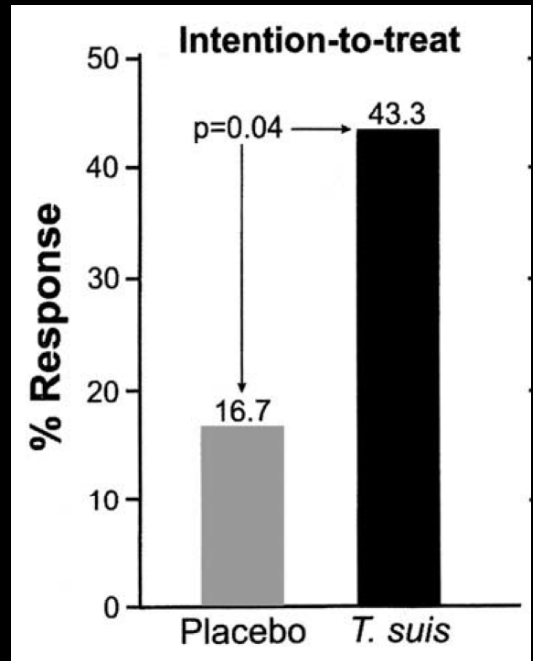
## CNDO-109 (Tumor-Activated Natural Killer Cells)

Acute Myeloid Leukemia	→				
Multiple Myeloma	→	2014 →	→		

\* TSO Phase 2a studies being conducted as Investigator-Initiated Studies; CNDO-109 Phase 2a AML study is a Phase 1/2 study

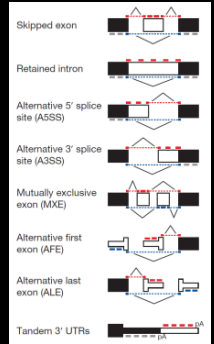
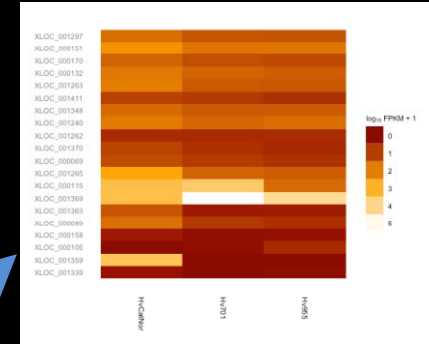
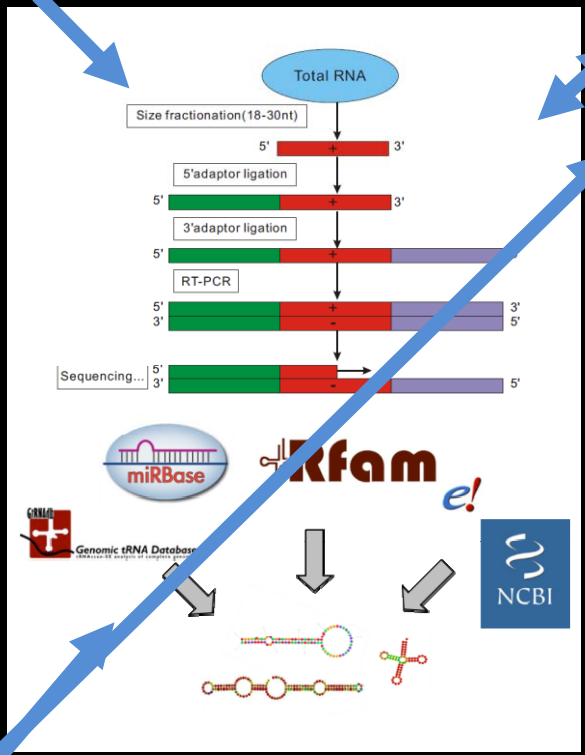
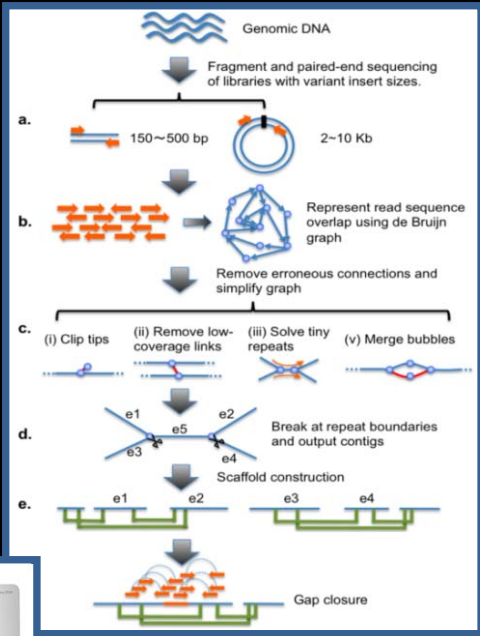
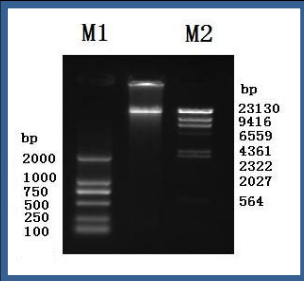
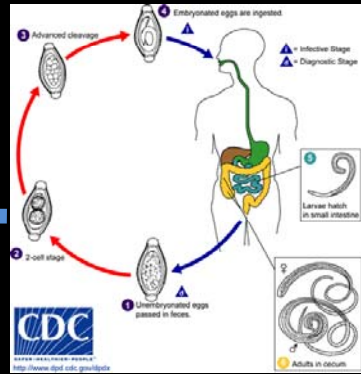
coronado  
BIOSCIENCES

Planned

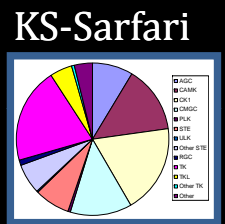
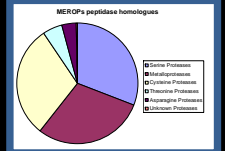
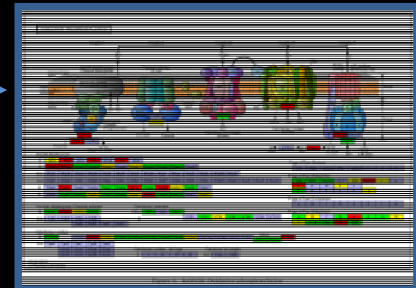




# Sampling, sequencing and analysis



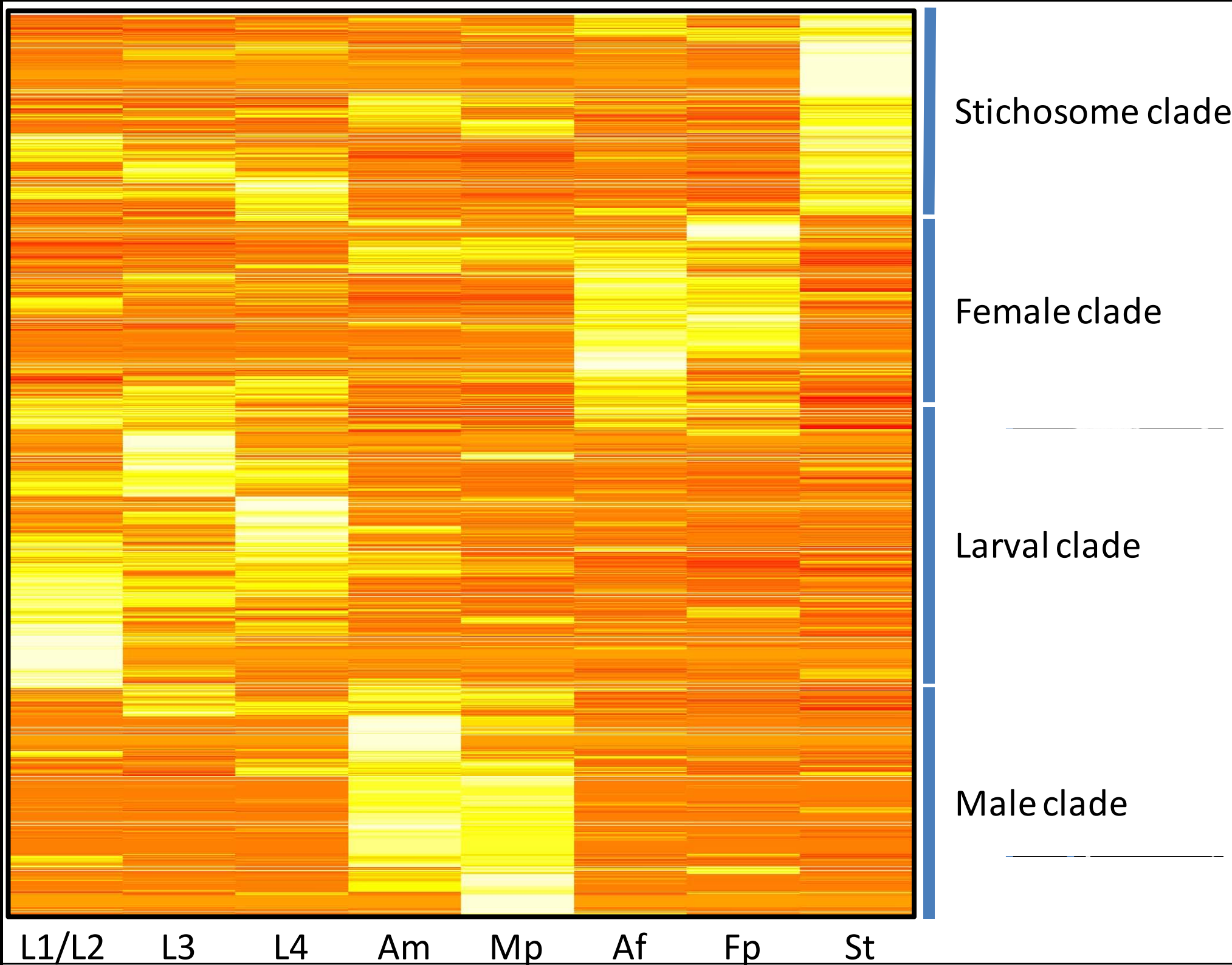
KEGG/KOBAS



# Male and female *T. suis* genome

**Table 1| Features of the *T. Suis* male and female genomes**

	<u>Male genome</u>	<u>Female genome</u>
Kmer (17) estimated genome size (Mb)	83.6	87.2
Total read data; estimated coverage	11.73 Gb; 140x	12.36 Gb; 142x
Total assembly size (in Mb; total contigs)	81.3; 60,856	76.0; 42,663
Total base pairs (Mb) in scaffolds > 200 bp; total number	74.2; 4,293	71.0; 3,288
Largest contig (Mb)	1.59	1.44
N50 in kb (contigs > 200 bp)	500	440
N90 in kb (contigs > 200 bp); total number > N90 length	81.0; 185	104; 168
% GC content (whole genome; contigs > 200 bp)	43.9; 43.6	43.6; 43.5
% Repetitive sequence (contigs > 200 bp)	31.7	32.3
% Coding (exonic; including introns)	28.9; 68.8	29.2; 64.4
Number of putative coding genes	14,540	14,395
Gene size (mean kb)	3.7	3.9
Exons per gene (mean)	5.4	5.7
Exon length (mean bp)	271	270
Introns per gene (mean)	4.4	4.7
Intron length (mean bp)	511	509
Number of transfer RNAs	991	1,021



Stichosome clade

Female clade

Larval clade

Male clade

L1/L2

L3

L4

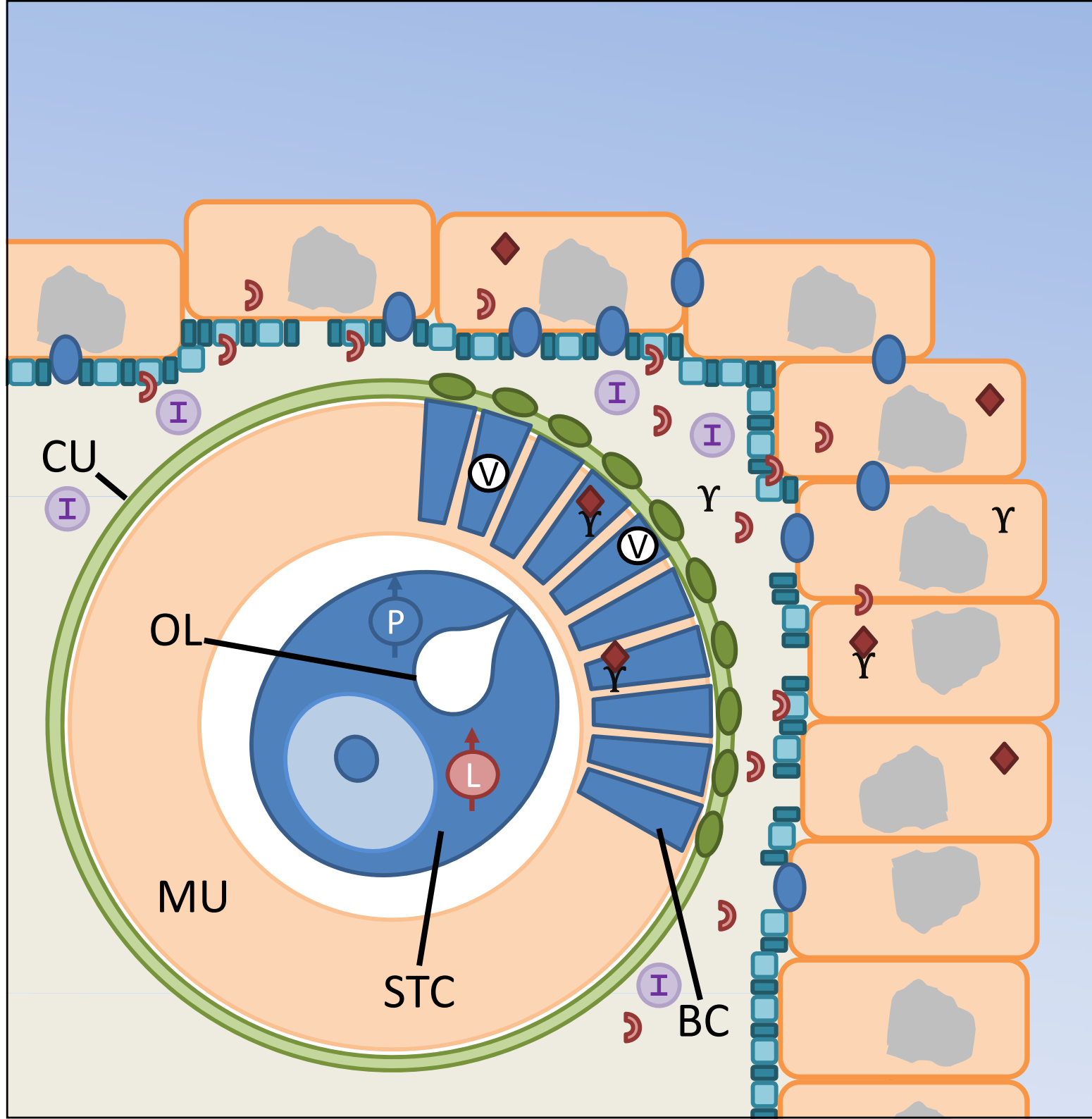
Am

Mp

Af

Fp

St



### ES products

- I Immunomodulators
- C Secreted proteases
- Y Small molecule transporters
- O Porins

### Metabolism

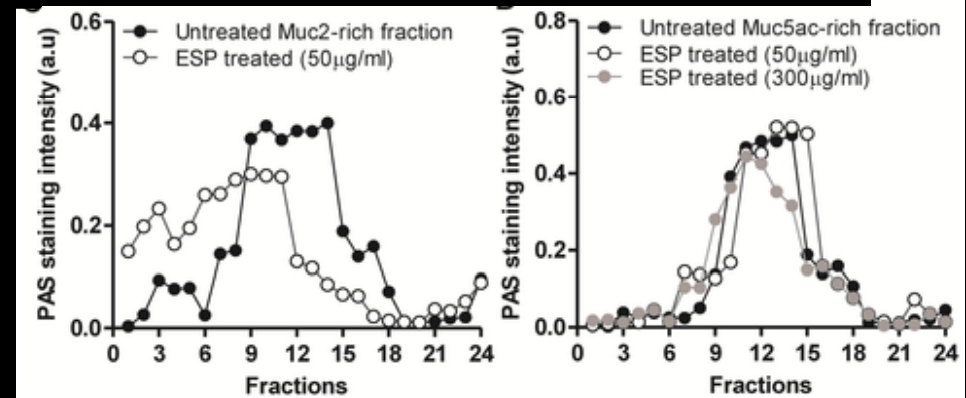
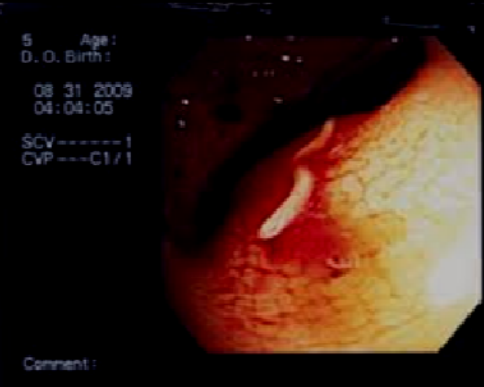
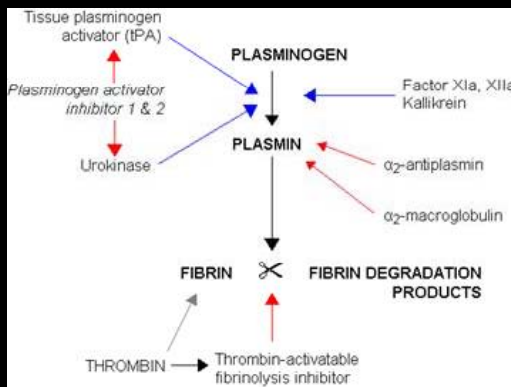
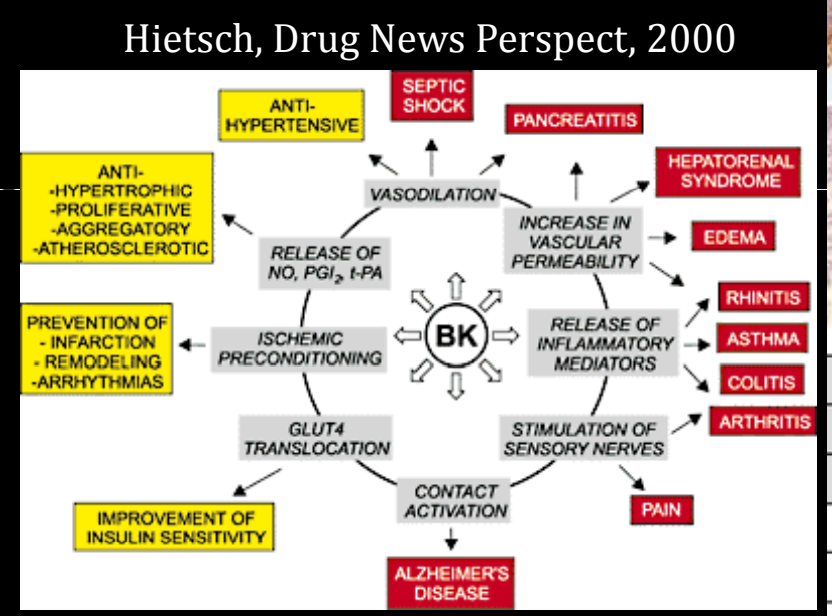
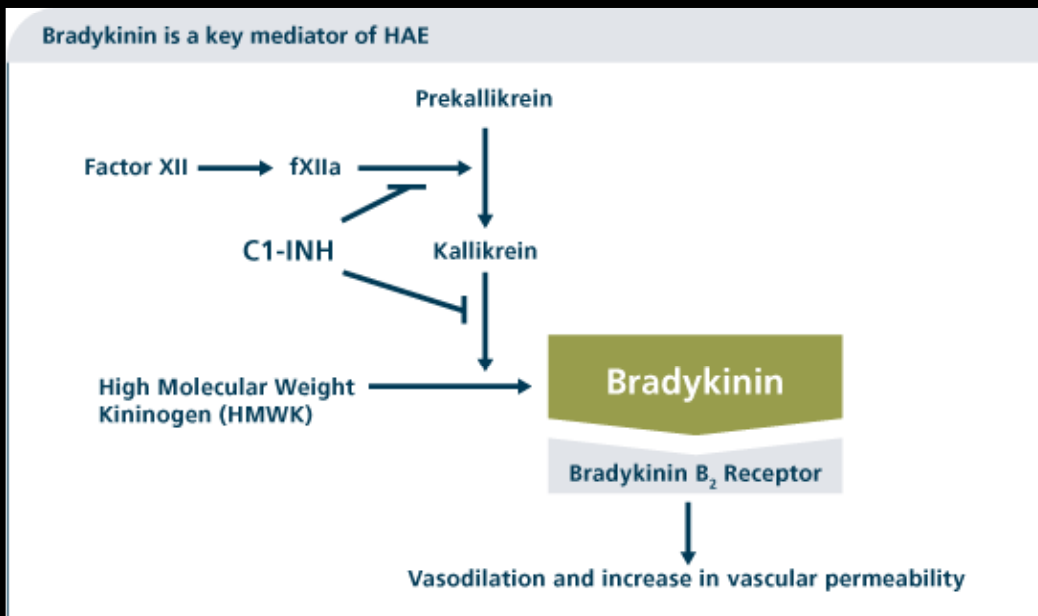
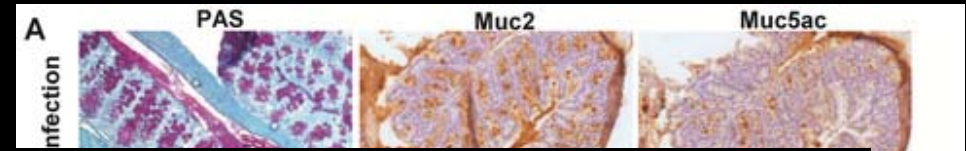
- P Peroxisome
- L Lysosome
- V Endocytic vesicle

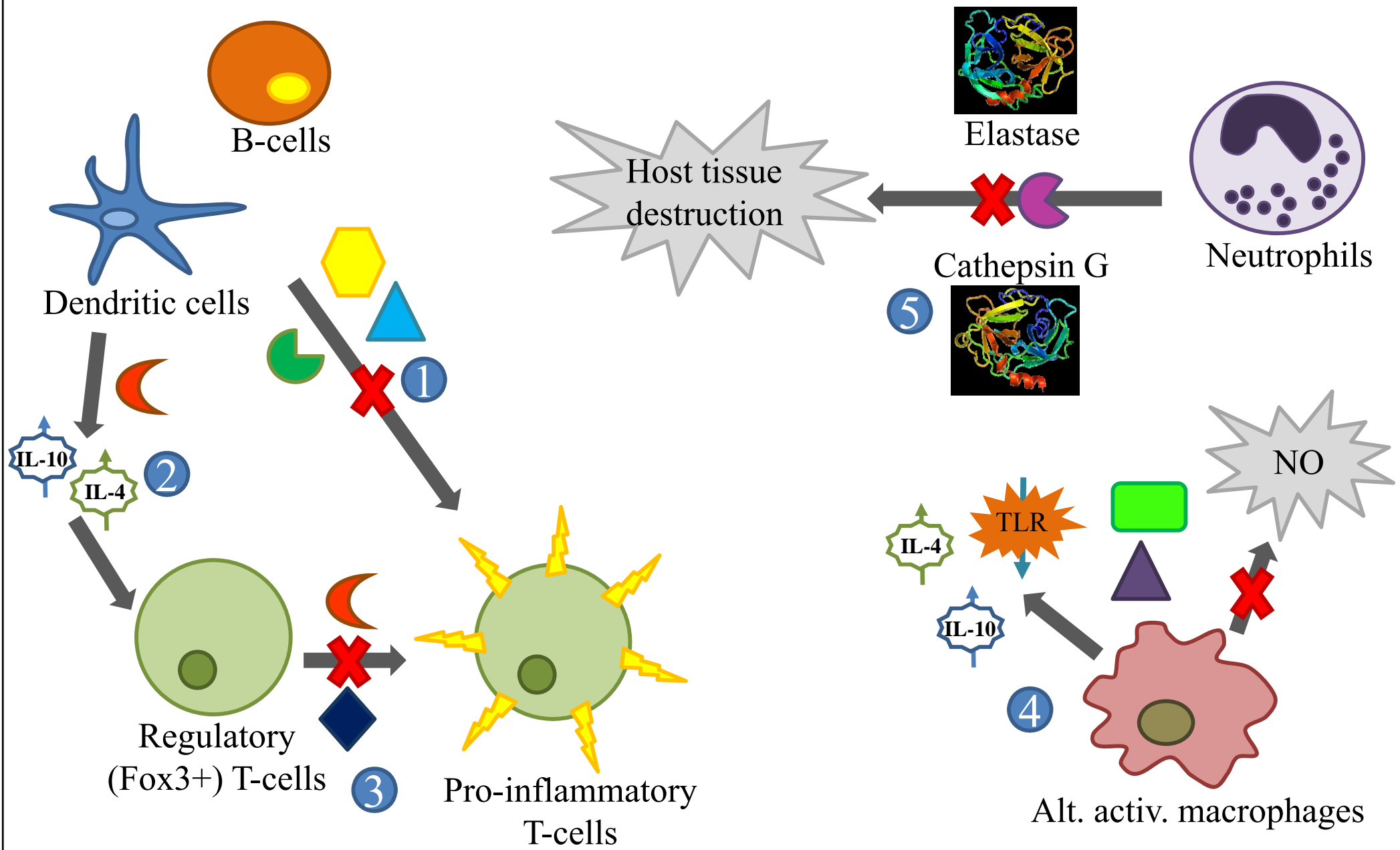
### Host molecules

- ◆ Small molecules
- Muc2
- Muc5α

# S1 (trypsin-like) serine proteases

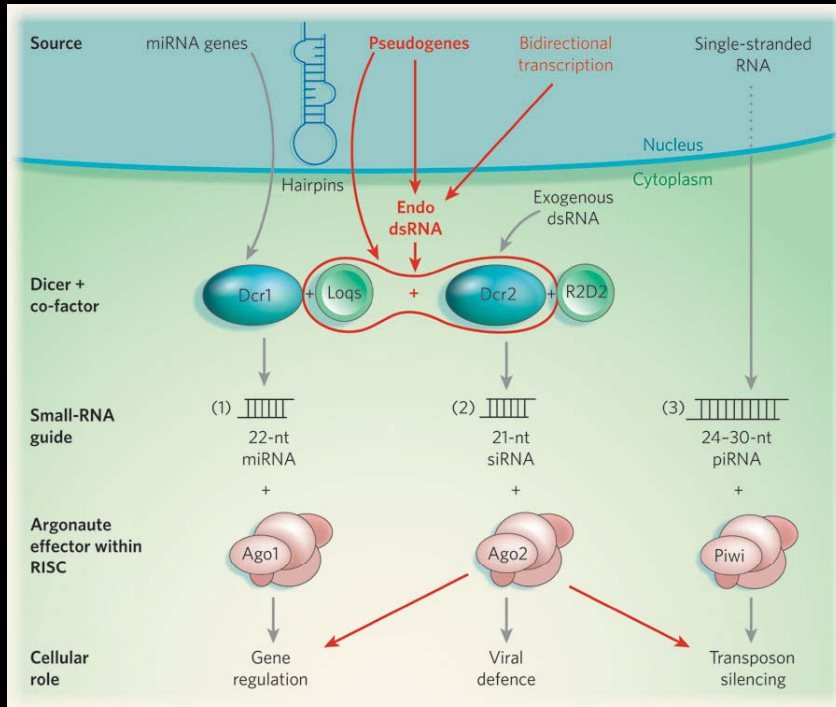
Hasnain et al . PLoS NTD, 2012





- |  |                  |   |              |   |       |   |                 |
|--|------------------|---|--------------|---|-------|---|-----------------|
|  | Serpin (TsCEI)   |  | Apyrase      |  | MIF   |  | TPX             |
|  | Cystatin (CPI-2) |  | Calreticulin |  | TGF-β |  | N-linked glycan |

# Non-coding RNAs



RNAi pathway complete

450 million reads

23,013 small RNA clusters

2 million consensus small RNAs

siRNAs - 22G/26G uncommon

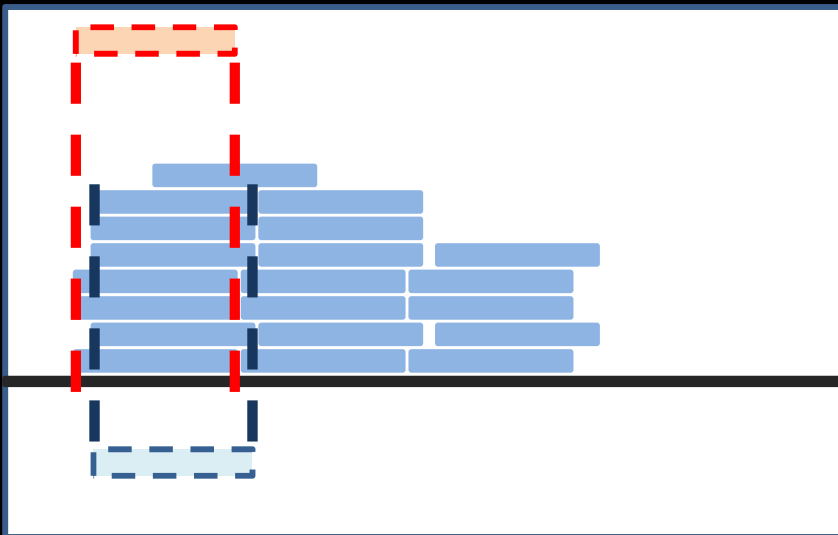
3,500 siRNA regulated genes

piwiRNAs - 21U-RNAs uncommon

- 200,000 anti-sense to TE

- role in germline?

miRNAs - 319 predicted

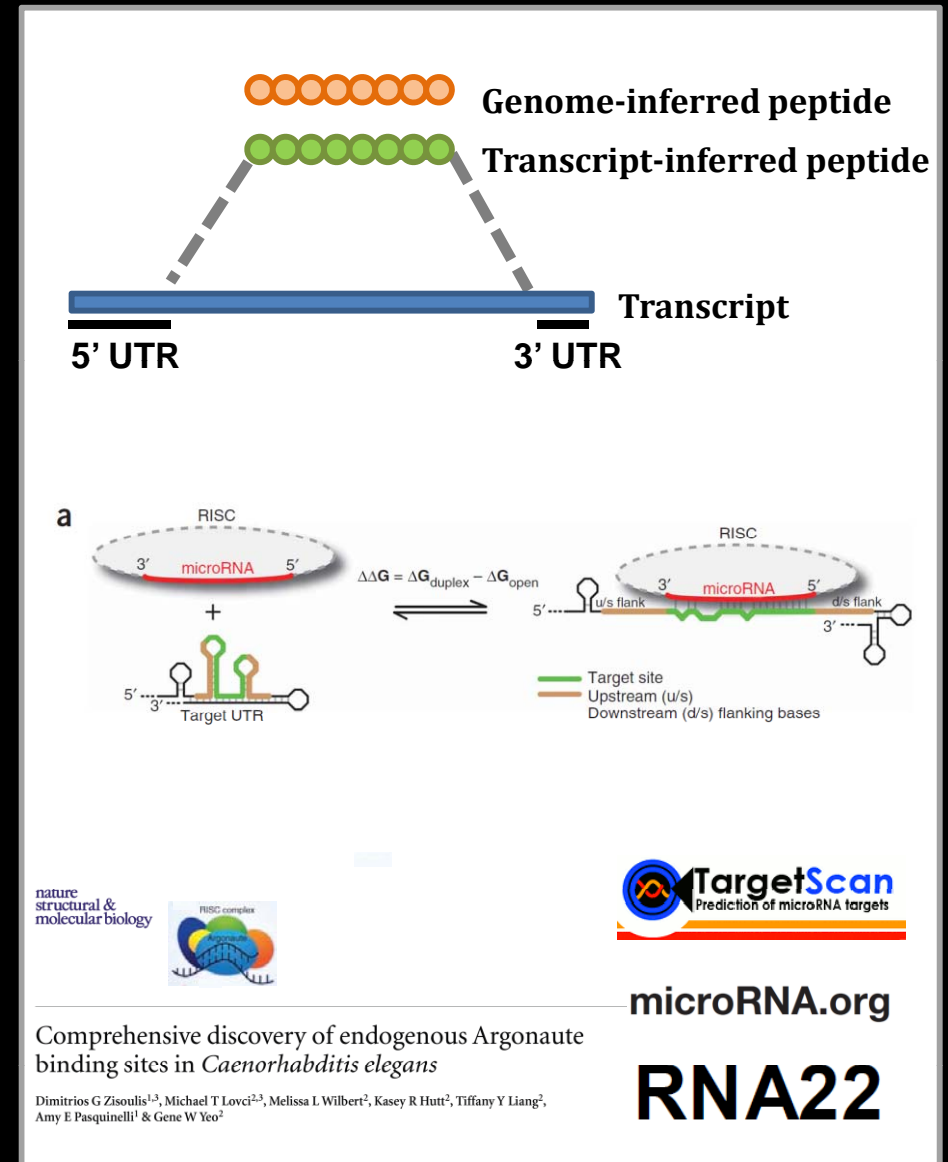


# miRNA:mRNA interactions

300,042 supported interactions

45 (62) miRNAs and 3,205 (3,877) genes

miR-61, miR-228, miR-236 and miR-252 key (maturation / sexual-diff.)





# miRNA:mRNA interactions

L4 to adulthood: miR-61 / miR-228 down

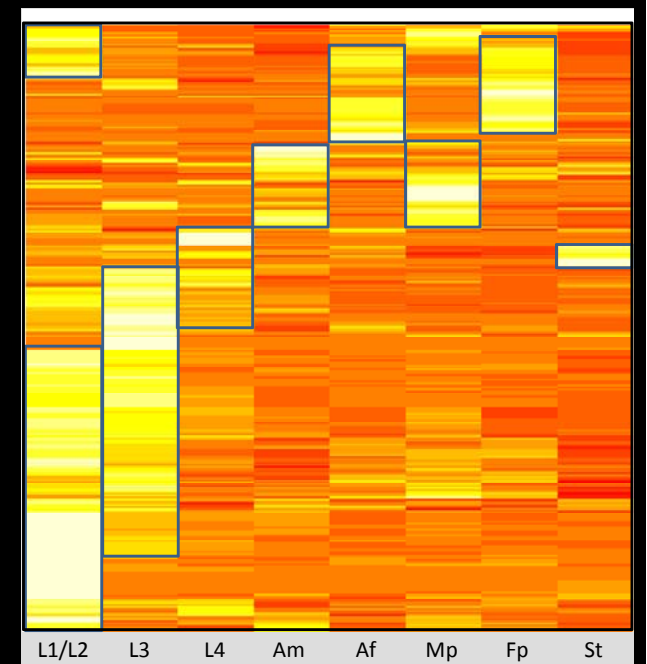
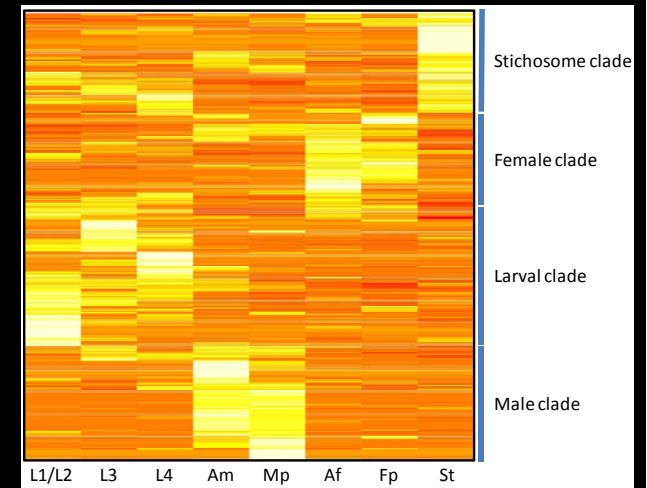
- **↑** 447 'co-regulated' genes in adults
- embryonic/genital development, reproduction, growth, metabolism (e.g., *tbx-2*, *vps-16* and *dyc-1*)

Male-enriched: miR228 up

- **↓** 320 'regulated' genes in males
- vulval/germline development, egg production, embryogenesis (e.g., *exc-4*, *unc-29*, *-58*, *bcat-1*)

Female-enriched: miR-236 / miR-252 up

- **↓** 205 'regulated' genes in females
- spermatogenesis, male mating/development, chemosensation (e.g., *cogc-5*, *goa-1* and *grk-2*)

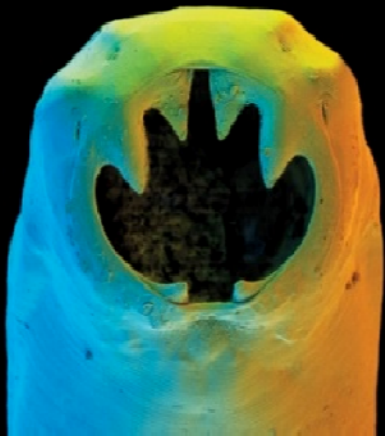


# Other STH genome projects

*Necator americanus* - WUSTL (nearly complete)

*Ancylostoma caninum* - SANGER (under way)

*Strongyloides* spp. - SANGER (under way)



# Acknowledgements

## Funding and support from:

- Australian Research Council Linkage
- National Health and Medical Research Council
- Beijing Genomics Institute
- Victorian Life-Sciences Initiative
- Danish Agency for Science, Technology and Innovation