Transcriptomic analysis of *Ascaris suum* larvae during their hepatopulmonary migration.

Wang T.[1] & Vlaminck J. [1], Jex A. [2], Gasser R. [2], Geldhof P. [1]

[1]*Faculty of veterinary medicine, Department of virology, parasitology & immunology. Salisburylaan 133, 9820 Merelbeke, Belgium.*
[2] Faculty of Veterinary Science, The University of Melbourne, Australia .
[*] *Corresponding Author: Johnny.vlaminck@ugent.be*

Introduction

Infections with the intestinal parasitic nematode *Ascaris sp.* present a problem in pig and human hosts all over the world. During the initial stage of infection, the invading stage three larvae (L3) pass through the intestinal tissue, liver and lungs to eventually grow into adulthood in the small intestine of its host (See Fig 1). During this migration, larvae have to continuously adapt to the changing environmental conditions as well as prevent possible fatal attack by the host's immune defences. All of this is likely to require tightly regulated transcriptional changes in the parasite.

Aim of the study

The intent of this project was to offer researchers an extra tool to investigate parasite biology, host-parasite interaction processes, and potential novel drug- or vaccine targets by producing the complete transcriptome of the *Ascaris suum* larvae during the different stages of their hepatotracheal migration using next generation sequencing.

Results

L3 liver

Most abundant transcripts

	# Contig	NCBI blast result
1	GS_16238	Cuticle collagen 12 [A. suum]
2	GS_13304	Cuticle collagen 155, partial [A. suum]
3	GS_18402	Cuticle collagen dpy-5, partial [A. suum]
4	GS_16291	Translation elongation factor 1A [A. suum]
5	GS_08970	Cuticle collagen 19, partial [A. suum]
6	GS_12737	Cuticle collagen 13, partial [A. suum]
7	GS_04352	Cuticle collagen 19, partial [A. suum]
8	GS_06634	ADP/ATP translocase [A. suum]
9	GS_14359	60S ribosomal protein L8 [A. suum]
10	GS_15085	Myosin regulatory light chain 1 [A. suum]
11	GS_20130	GAPDH [A. suum]
12	GS_07034	troponin-like protein [Anisakis simplex]
13	GS_00138	Myoglobin [A. suum]
14	GS_14911	GAPDH, partial [A. suum]
15	GS_17425	GAPDH, partial [A. suum]
16	GS_20320	40S ribosomal protein S12 [A. suum]
17	GS_20031	Actin-2, partial [A. suum]
18	GS_10291	Actin-2, partial [A. suum]
19	GS_07489	Fatty-acid and retinol-binding protein 1 [A. suum]
20	GS_14475	60S ribosomal protein L18a, partial [A. suum]

Stage specific transcripts

	# Contig	NCBI blast result	
1	GS_08770	Unknown	
2	GS_03315	Unknown	
3	GS_23463	Unknown	
4	GS_08624	Neprilysin-1 [A. suum]	
5	GS_07157	Esterase [A. suum]	
6	GS_02759	Galactoside 2-alpha-L-fucosyltransferase [A.suum]	
7	GS_20830	Unknown	
8	GS_13954	Cell division control protein 42 [A. suum]	
9	GS_13462	Nuclear pore complex protein 15 [A. suum]	
10	GS_09097	Unknown	
11	GS_22364	Serine/threonine-protein phosphatase [A. suum]	
12	GS_01361	Unknown	
13	GS_20942	Tyrosine-protein phosp. non-receptor [A. suum]	
14	GS_06723	Protein SPEC3 [A. suum]	
15	GS_00277	Dynein light chain 2 [A. suum]	
16	GS_01069	Integrator complex subunit 4 [A. suum]	
17	GS_09830	Hyaluronidase-1, partial [A. suum]	←
18	GS_21471	Unknown	
19	GS_17399	Integrator complex subunit 3 [A. suum]	
20	GS_05490	Chondroitin proteoglycan-2 [A. suum]	

L3 egg

Most abundant transcripts

	# Contig	NCBI blast result
1	GS_16291	Translation elongation factor 1A [A. suum]
2	GS_14359	60S ribosomal protein L8 [A. suum]
3	GS_06634	ADP/ATP translocase [A. suum]
4	GS_15457	40S ribosomal protein S20 [A. suum]
5	GS_04971	Cleavage stimulation factor subunit 2 [A. suum]
6	GS_14905	Tubulin alpha chain [A. suum]
7	GS_07536	40S ribosomal protein S30 [A. suum]
8	GS_20130	GAPDH [A. suum]
9	GS_03797	heat shock protein 70 [Dracunculus medinensis]
10	GS_16337	40S ribosomal protein S5, partial [A. suum]
11	GS_15291	60S ribosomal protein L24 [A. suum]
12	GS_19340	60S ribosomal protein L27a [A. suum]
13	GS_14911	GAPDH, partial [A. suum]
14	GS_17425	GAPDH, partial [A. suum]
15	GS_01968	40S ribosomal protein S15 [A. suum]
16	GS_23993	Tubulin beta-1 chain, partial [A. suum]
17	GS_10985	40S ribosomal protein S9 [A. suum]
18	GS_12185	Neuroendocrine protein 7B2 [A. suum]
19	GS_21140	Elongation factor 2, partial [A. suum]
20	GS_07034	Troponin C, isoform 2, partial [A. suum]

Stage specific transcripts

	# Contig	NCBI blast result	
1	GS_10173	Unknown	
2	GS_20139	Unknown	
3	L3E_00423	Unknown	
4	GS_00826	Venom allergen 5 [A. suum]	←
5	GS_18821	Peroxisomal biogenesis factor 3 [A. suum]	
6	L3E_01194	Unknown	
7	L3E_00108	Unknown	
8	GS_16703	Unknown	
9	GS_21056	RAG1-activating protein 1 [A. suum]	
10	GS_22077	catalase [A. suum]	
11	L3E_03523	Transcription initiation factor [A. suum]	
12	L3E_00774	Unknown	
13	GS_00224	Unknown	
14	L3E_00619	Unknown	
15	L3E_01939	Unknown	
16	L3E_01111	Venom allergen 5 [A. suum]	←
17	GS_24358	Plexin-2 [A. suum]	
18	GS_11924	Unknown	
19	GS_19731	Cyclin-G-associated kinase [A. suum]	
20	GS_19471	Unknown	

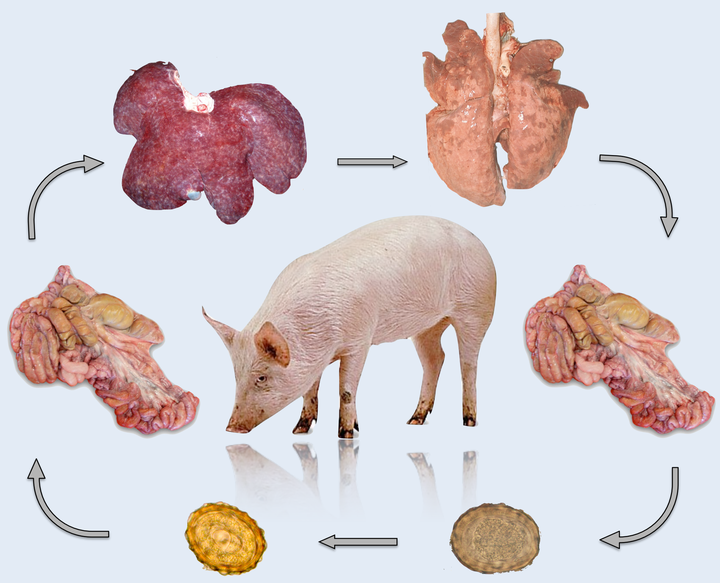

Fig. 1: Life cycle of *Ascaris suum*. Infective stage 3 larvae (L3 egg) hatch from ingested *Ascaris* eggs, penetrate the intestinal wall and migrate through the liver (L3 liver) and the lungs (L3 lung). Larvae crawl up the respiratory tree, get swallowed again and grow into adulthood in the small intestine (L4, L5 & Adults).

L3 lung

Most abundant transcripts

	# Contig	NCBI blast result	
1	Liv_03570	Unknown	←
2	GS_17199	C-type lectin Tc-ctl-4 [Toxocara canis]	←
3	GS_16238	Cuticle Collagen 12 [A. suum]	
4	GS_22776	excretory/secretory mucin MUC-5 [Tox. canis]	←
5	GS_00138	Myoglobin [A. suum]	
6	GS_08970	Cuticle collagen 19 [A. suum]	
7	GS_18402	Cuticle collagen dpy-5, partial [A. suum]	
8	GS_16291	eukaryotic translation elongation factor [A. suum]	
9	GS_12737	Cuticle collagen 13, partial [A. suum]	
10	GS_07489	Fatty-acid and retinol-binding protein [A. suum]	
11	GS_08505	Troponin T [A. suum]	
12	GS_16792	Cuticle collagen 14 [A. suum]	
13	GS_04352	Cuticle collagen 19 [A. suum]	
14	GS_06453	As16 [A. suum]	←
15	GS_24170	Electron transfer flavoprotein subunit β [A. suum]	
16	GS_13304	Cuticle collagen 155, partial [A. suum]	
17	GS_21096	Peroxiredoxin, partial [A. suum]	←
18	GS_06634	ADP/ATP translocase [A. suum]	
19	GS_19340	60S ribosomal protein L27a [A. suum]	
20	GS_15085	Myosin regulatory light chain 1 [A. suum]	

Stage specific transcripts

	# Contig	NCBI blast result	
1	Lung_01340	Unknown	
2	GS_14656	Unknown	
3	GS_10377	Unknown	
4	GS_00710	Unknown	
5	GS_22691	Thioredoxin-1, partial [A. suum]	←
6	GS_11637	Unknown	
7	Liv_01875	Unknown	
8	Lung_00162	Unknown	
9	GS_24217	Prolyl 4-hydroxylase subunit alpha-1 [A. suum]	
10	GS_16512	LIM and calponiny domain-cont. prot. [A. suum]	
11	Lung_01110	Unknown	
12	GS_00999	PHD finger & CXXC domain-cont. prot. [A. suum]	
13	GS_09081	Homeobox protein ceh-20 [A. suum]	
14	GS_20551	Unknown	
15	GS_08905	Unknown	
16	GS_24301	Phosphatidylinositol transfer prot. β [A. suum]	
17	GS_05175	Unknown	
18	GS_21771	E3 ubiquitin-protein ligase UBR5 [A. suum]	
19	L3E_02788	Unknown	
20	GS_17452	Anion exchange protein 3 [A. suum]	

L4

Most abundant transcripts

	# Contig	NCBI blast result	
1	GS_10576	Lymphocyte antigen 75 [A. suum]	←
2	GS_08970	Cuticle collagen 19 [A. suum]	
3	GS_05876	Chymotrypsin/elastase isoinhibitor 1 [A. suum]	
4	GS_16291	eukaryotic translation elongation factor [A. suum]	
5	L4_01560	similar to astacin-like protein	←
6	GS_02102	As14 [A. suum]	←
7	GS_18402	Cuticle collagen dpy-5, partial [A. suum]	
8	GS_07489	Fatty-acid and retinol-binding protein 1 [A. suum]	
9	GS_12783	Cuticle collagen 12, partial [A. suum]	
10	GS_20130	GAPDH [A. suum]	
11	GS_06634	ADP/ATP translocase [A. suum]	
12	GS_15085	Myosin regulatory light chain 1 [A. suum]	
13	GS_14911	GAPDH, partial [A. suum]	
14	GS_17425	GAPDH, partial [A. suum]	
15	GS_19276	Fructose-bisphosphate aldolase 1 [A. suum]	
16	GS_06453	As16 [A. suum]	←
17	GS_10438	Acetyl-CoA acetyltransferase A [A. suum]	
18	GS_20031	Actin-2, partial [A. suum]	
19	GS_04841	Actin-2, partial [A. suum]	
20	GS_12737	Cuticle collagen 13, partial [A. suum]	

Stage specific transcripts

	# Contig	NCBI blast result	
1	GS_03310	Unknown	
2	GS_01929	Unknown	
3	GS_03875	Unknown	
4	GS_06595	CRAL-TRIO domain-containing protein [A.suum]	
5	GS_03349	Cuticle collagen 12 [A. suum]	
6	GS_13199	Unknown	
7	GS_10842	Unknown	
8	GS_14755	Nematode cuticle collagen protein [B. malayi]	
9	GS_16132	Cuticle collagen 12, partial [A. suum]	
10	GS_22474	Unknown	
11	GS_18539	Unknown	
12	GS_15741	Trehalase [A. suum]	
13	GS_23638	Nematode cuticle collagen protein [B. malayi]	
14	GS_06967	Cecropin-P2 [A. suum]	←
15	GS_04298	Unknown	
16	GS_20992	acyltransferase [Loa loa]	
17	GS_21696	Exocyst complex component 5 [A. suum]	
18	GS_07406	Integrin beta pat-3 [A. suum]	
19	L4_03487	Hypothetical protein CBG19295 [C. briggsae]	
20	GS_02308	Unknown	

Cuticle collagen expression during the hepatotracheal migration of *A. suum* larvae

contig	RPKM L3 egg	RPKM L3 liver	RPKM L3 lung	RPKM L4	*C. elegans* gene	*C. elegans* % identity
GS_16238	1534,1	41285,9	31329,1	1401,5	col-73	49%
GS_12783	28,4	1040,3	362,1	9335,6	col-73	64%
GS_13798	0,1	9,4	4,3	495,2	col-73	72%
GS_11610	63,7	738,5	253,0	1708,5	col-73	35%
GS_16132	0,1	0,1	0,5	796,6	col-68	43%
GS_12737	466,0	12333,3	11825,1	5333,0	col-68	62%
GS_14767	12,8	794,2	515,4	2772,3	col-166	72%
GS_03349	0,1	0,4	0,8	989,4	col-166	72%
GS_16030	169,4	430,6	565,1	515,4	col-167	75%
GS_16792	92,3	5204,6	8397,4	1252,8	col-48	46%
GS_15987	169,2	4006,2	3122,1	1940,9	col-14	46%
GS_09547	167,6	1790,5	1965,3	2401,4	col-3	78%
GS_01804	79,7	3880,7	4985,6	1162,7	col-109	61%
GS_22906	159,6	3864,7	4315,0	2098,0	col-97	53%
GS_00618	40,7	3263,4	4309,4	2325,0	col-104	35%
GS_15592	160,2	2556,6	1363,0	1114,8	col-34	57%
GS_13304	1488,2	22411,6	7754,1	31,0	col-154	23%
GS_08765	361,4	3939,8	5866,4	1503,6	col-182	37%
L4_00638	1,8	395,4	230,8	2261,8	col-184	45%
GS_18402	635,2	17849,0	16707,4	10912,4	col-184	44%
GS_23251	245,9	4600,6	2487,8	1189,5	col-184	44%
GS_08970	708,2	12666,0	18270,7	13752,6	col-184	44%
GS_02306	6,8	0,1	0,1	468,0	col-130	42%
GS_18250	4,7	645,1	915,9	2013,2	col-65	31%
GS_04352	544,5	11483,7	8151,2	2284,5	col-65	55%
GS_21063	0,6	6,0	263,5	2081,7	col-35	64%
GS_14343	0,4	1,5	18,5	509,9	col-150	56%
GS_12116	18,2	1029,5	2492,6	3990,6	col-150	58%
GS_21873	0,1	335,7	928,9	1281,2	col-63	50%
GS_08501	348,8	411,2	461,9	561,7	α2 type IV col	86%

Fig. 2: Phylogenetic relation and expression levels of the different cuticle collagen transcripts detected in the top 1000 most produced transcripts of the L3 egg, L3 liver, L3 lung and L4 of *Ascaris suum*. (RPKM = Reads Per Kilobase per Million reads).

The Top-100 most abundant transcripts per life stage were investigated. Results showed that in the egg L3, transcripts associated with the regulation of translation and transcription, mainly ribosomal proteins, were most abundant. Next to that, several other interesting transcripts were found (see black arrows). Namely, the protective antigens As14 & As16, the antioxidant peroxiredoxin, the astacin-like metalloprotease and the C-type lectin-4, lymphocyte antigen 75 or secretory mucin 5 which might have important functions in host mimicking and/or immune evasion.

From the liver-L3s onwards, high transcription levels were seen for cuticle collagens, indicating the growth of the larvae during their migration. The expression profile of all collagen genes present in the Top 1000 transcripts is represented in figure 2.

Transcripts were also sorted according to their life stage specificity by sorting the transcripts on descending RPKM value and subsequently filtering out the transcripts with at least a 100-fold higher RPKM value in comparison with other life stages. Interesting transcripts were the VAP-5 antigen in the egg L3, the hyaluronidase in the larvae penetrating the liver, the antioxidant thioredoxin in the lung L3 and the anti-bacterial substance Cecropin 2B in the L4. A great number of transcripts did not show any homology to other proteins within the NCBI database (e-value cut-off = 10^{-5}) implying that many biologically interesting molecules from this parasite remain to be discovered and investigated.

Acknowledgements

This work was financed by the Flemish Agency for Innovation through Science and Technology.