

NORNEX Consortium

***Fungal genomics strand:
The genome biology of
Hymenoscyphus pseudoalbidus***

Original aims:

[from proposal]

To generate genome sequences of up to 30 isolates of *C. fraxinea* from the UK and Europe, and multiple isolates of the related non-pathogenic fungus *H. albidus*.

Genome sequences will

- reveal the origins of the pathogen,
- provide markers to allow the spread of different strains to be followed
- identify whether the rapid spread is due to crosses with *H. albidus*
- identify potential virulence factors such as secreted effectors that may be crucial to the invasive nature of the pathogen
- underpin interpretation of transcriptomic data

& Comparison of UK isolates to isolates from continental Europe at a whole-genome level will be highly informative.

& Complement existing sequencing efforts by our collaborators in Europe.

Revised aims

- What is the diversity of *H. pseudoalbidus* in Europe?
- Was there one introduction or many?
- Are strains interbreeding?
- Are there patterns of diversity with time of isolation in European *H. pseudoalbidus*?
- Is there association of genomic diversity with geography (and thus perhaps ash host genotype)?
- What was/were the origin(s) of the UK infections?

AGENDA: *H. pseudoalbidus*

Matt Clark

KI reference genome sequencing

Adam Vivian-Smith

Tracking diversity in Norway with ddRAD

Mark Blaxter

Preliminary genome sequencing: 23 strains

NORNEX EDINBURGH

Georgios Koutsovoulos



***H. pseudoalbidus* from Europe**

New genome data

SAMPLES FROM

Adam Vivian-Smith & Ari Hietala

Skog og Landskap - Norwegian Forest and Landscape Institute

Bioforsk - Norwegian Institute for Agricultural and Environmental Research

Renaud Ios & colleagues

Unité de Mycologie, Laboratoire de la Santé des Végétaux, Domaine de
Pixérécourt, Malzéville, France

***H. pseudoalbidus* from Europe**

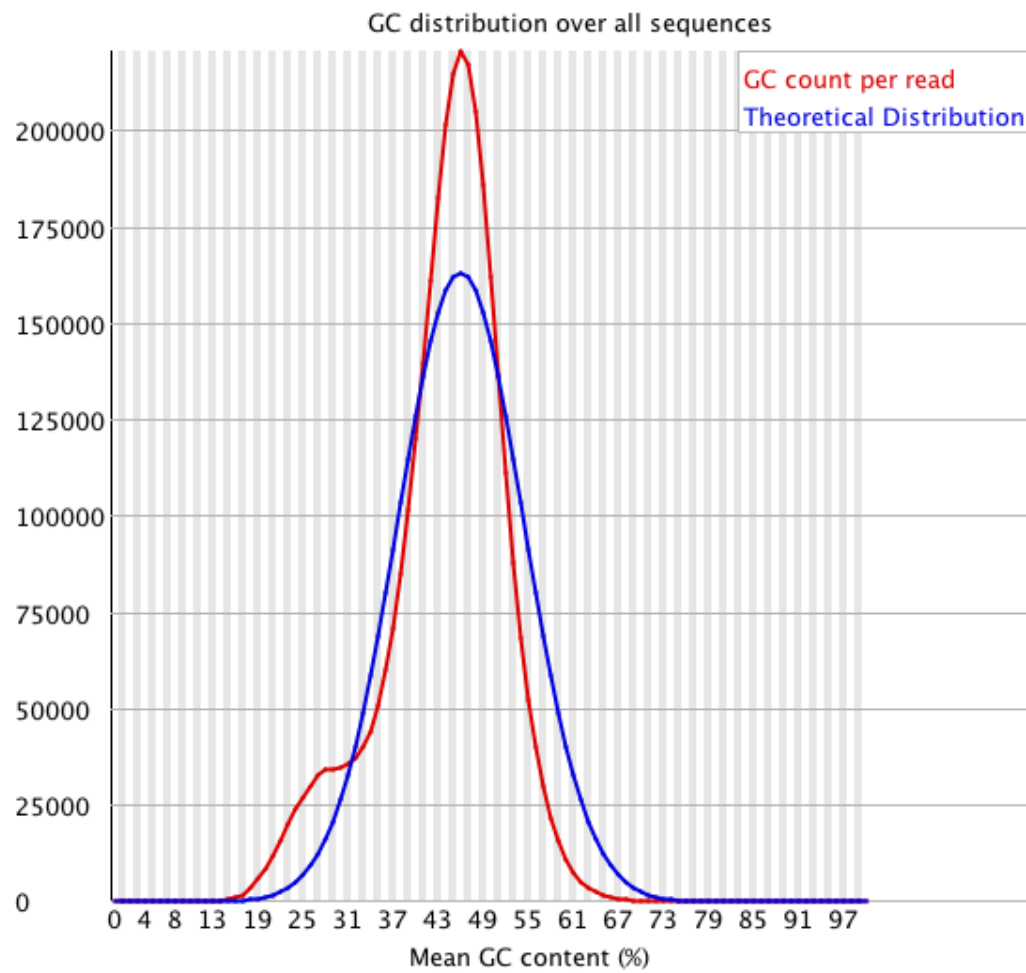
NORNEX ID	YEAR	COUNTRY
H_2008_125_2	2008	NORWAY
H_2008_139_1	2008	NORWAY
H_2008_142_5	2008	NORWAY
H_2008_148_4	2008	NORWAY
H_2008_152_4	2008	NORWAY
H_2008_81_6	2008	NORWAY
H_2009_86_3	2009	NORWAY
H_2010_189_4	2010	NORWAY
H_2010_189_5	2010	NORWAY
H_2011_11_1	2011	NORWAY
H_2012_24_1	2012	NORWAY
H_2012_38_2_2	2012	NORWAY
H_2012_42_1_1	2012	NORWAY
H_CBS122191	?	AUSTRIA
H_CBS122503	2011	POLAND
H_CBS122504	2005	POLAND
H_CBS122505	2000	POLAND
H_CBS122507	2000	POLAND
H_FON_M_1	2009	FRANCE
H_GIR_M_2	2009	FRANCE
H_LAN_M_1	2009	FRANCE
H_MIG_M_1	2009	FRANCE
H_LSTM82	2008	FRANCE

H. pseudoalbidus from Europe

NORNEX ID	Number of reads (million)	Span of data (gigabases)	Fold coverage
MIG-M-1	15.04	2.256	40.9
LAN-M-1	17.38	2.607	48.0
FON-M-1	18.09	2.713	49.0
CBS122191	18.73	2.809	50.7
CBS122507	13.54	2.031	37.1
CBS122503	14.36	2.154	39.3
CBS122505	14.11	2.117	38.5
2008-81-6	10.73	1.610	30.3
2008-148-4	10.21	1.532	28.7
2008-125-2	15.49	2.324	42.2
2008-152-4	13.95	2.092	38.3
2008-139-1	17.42	2.612	47.2
2008-142-5	20.82	3.123	55.7
2009-86-3	17.01	2.551	46.9
2010-189-4	16.88	2.533	46.7
2010-189-5	19.32	2.898	53.3
2011-11-1	15.23	2.284	41.6
2012-24-1	18.16	2.724	49.4
2012-38-2-2	23.85	3.578	64.0
2012-42-1-1	15.21	2.282	41.6

H. pseudoalbidus from Europe

NORNEX ID	contigs >500	span of contigs	N50	GC%
H_2008_125_2	6267	53.5	24303	42.67
H_2008_139_1	6041	53.5	24414	42.69
H_2008_142_5	6138	53.5	25254	42.72
H_2008_148_4	7560	51.8	18776	43.18
H_2008_152_4	6736	52.8	22327	42.9
H_2008_81_6	7387	51.6	19538	43.19
H_2009_86_3	6160	53.5	25295	42.72
H_2010_189_4	6348	53.3	23729	42.78
H_2010_189_5	6331	53.4	24880	42.77
H_2011_11_1	6118	53.4	25094	42.73
H_2012_24_1	6317	53.3	24080	42.77
H_2012_38_2_2	6094	53.5	24711	42.67
H_2012_42_1_1	6516	53.1	23078	42.83
H_CBS122191	6252	53.5	24172	42.71
H_CBS122503	6700	53.2	22119	42.77
H_CBS122504	6365	56.7	30182	41.76
H_CBS122505	6422	53.5	23708	42.67
H_CBS122507	6387	53.1	23038	42.79
H_FON_M_1	6436	53.3	23037	42.73
H_GIR_M_2	4770	51.9	34647	43.59
H_LAN_M_1	6534	52.5	23965	43.15
H_MIG_M_1	6498	52.8	23204	42.91
H_LSV82	35509	77.3	19479	38.33



Raw data

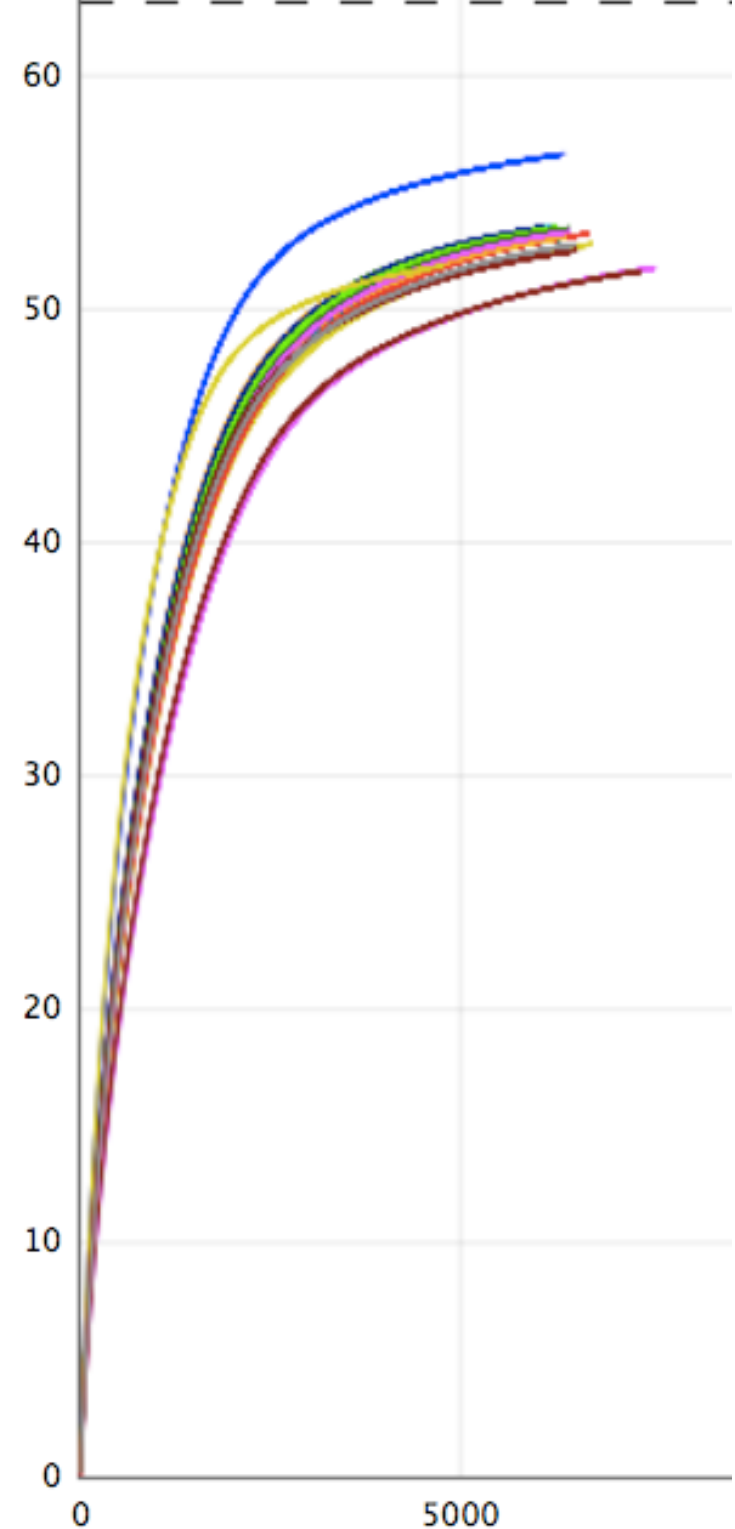
all have the “AT hump” noted previously

Initial assemblies

with CLC, single end QC assemblies

23 European strains; 35 to 100 fold coverage

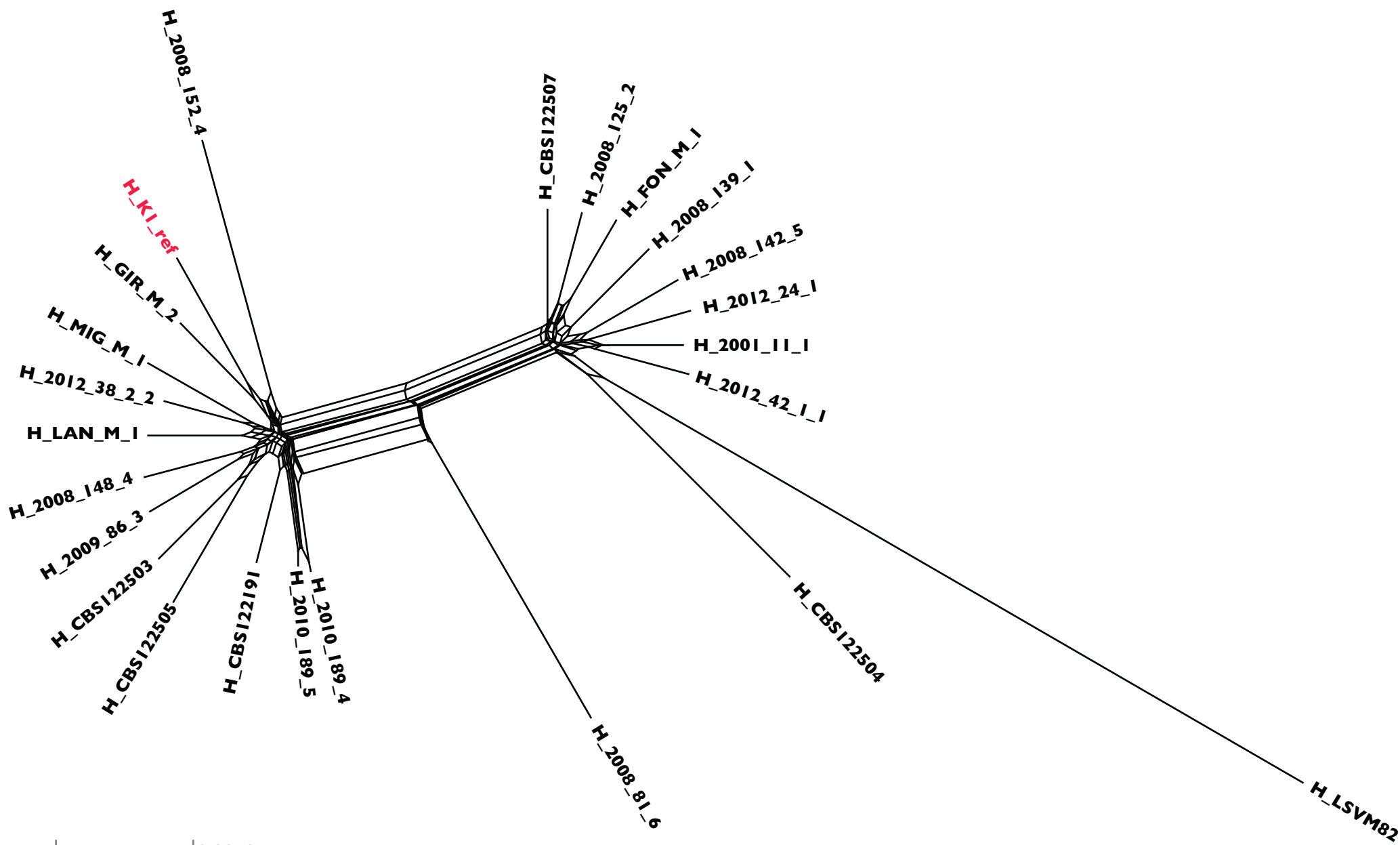
one strain (LSVM82) is odd (>>span, GC% different)



H. pseudoalbidus from Europe

NORNEX ID	SNP from KI	SNP / 100k	indels from KI	indels / 100k
H_2008_125_2	193829	400.56	37014	76.49
H_2008_139_1	174930	362.54	32599	67.56
H_2008_142_5	180469	376.75	32502	67.85
H_2008_148_4	175814	371.5	32875	69.47
H_2008_152_4	176916	367.83	34024	70.74
H_2008_81_6	187222	397.2	36522	77.48
H_2009_86_3	164775	336.36	32603	66.55
H_2010_189_4	179776	373.46	33438	69.46
H_2010_189_5	179106	371.87	33429	69.41
H_2011_11_1	188201	386.67	37279	76.59
H_2012_24_1	177522	367.95	33805	70.07
H_2012_38_2_2	174888	361.19	33466	69.12
H_2012_42_1_1	173764	358.91	32218	66.55
H_CBS122191	181217	376.99	32820	68.28
H_CBS122503	195226	408.27	35872	75.02
H_CBS122504	191934	372.98	33400	64.91
H_CBS122505	172436	355.22	32440	66.83
H_CBS122507	180121	373.09	34599	71.67
H_FON_M_1	179944	372.62	35230	72.95
H_GIR_M_2	196420	418.52	37771	80.48
H_LAN_M_1	168898	351.73	31707	66.03
H_MIG_M_1	171274	351.67	34758	71.37
H_LSTM82	200012	421.31	36898	77.72

Georgios Koutsovoulos

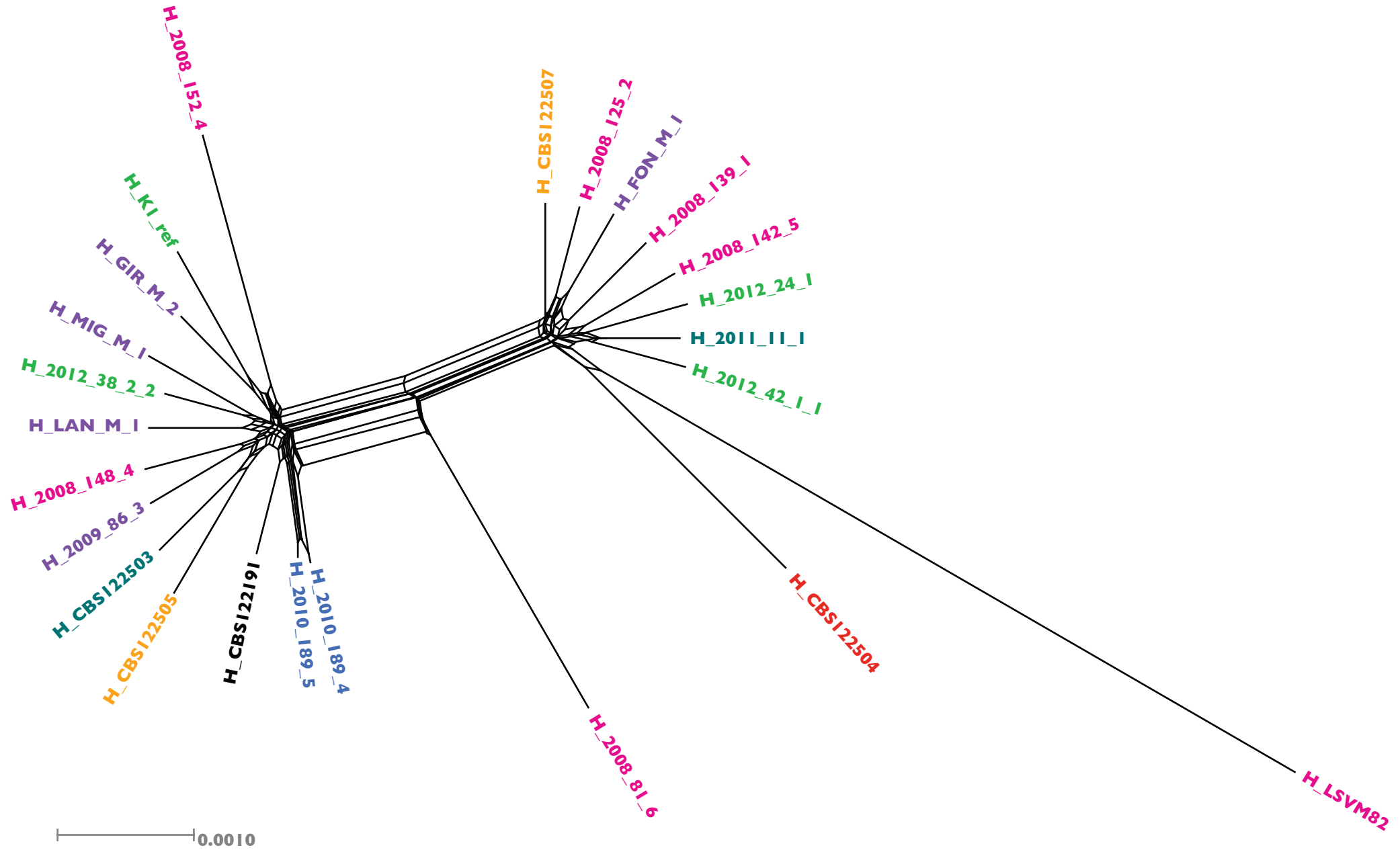


0.0010

***H. pseudoalbidus* from Europe**

NORNEX ID	YEAR	COUNTRY
H_2008_125_2	2008	NORWAY
H_2008_139_1	2008	NORWAY
H_2008_142_5	2008	NORWAY
H_2008_148_4	2008	NORWAY
H_2008_152_4	2008	NORWAY
H_2008_81_6	2008	NORWAY
H_2009_86_3	2009	NORWAY
H_2010_189_4	2010	NORWAY
H_2010_189_5	2010	NORWAY
H_2011_11_1	2011	NORWAY
H_2012_24_1	2012	NORWAY
H_2012_38_2_2	2012	NORWAY
H_2012_42_1_1	2012	NORWAY
H_CBS122191	?	AUSTRIA
H_CBS122503	2011	POLAND
H_CBS122504	2005	POLAND
H_CBS122505	2000	POLAND
H_CBS122507	2000	POLAND
H_FON_M_1	2009	FRANCE
H_GIR_M_2	2009	FRANCE
H_LAN_M_1	2009	FRANCE
H_MIG_M_1	2009	FRANCE
H_L SVM82	2008	FRANCE

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012

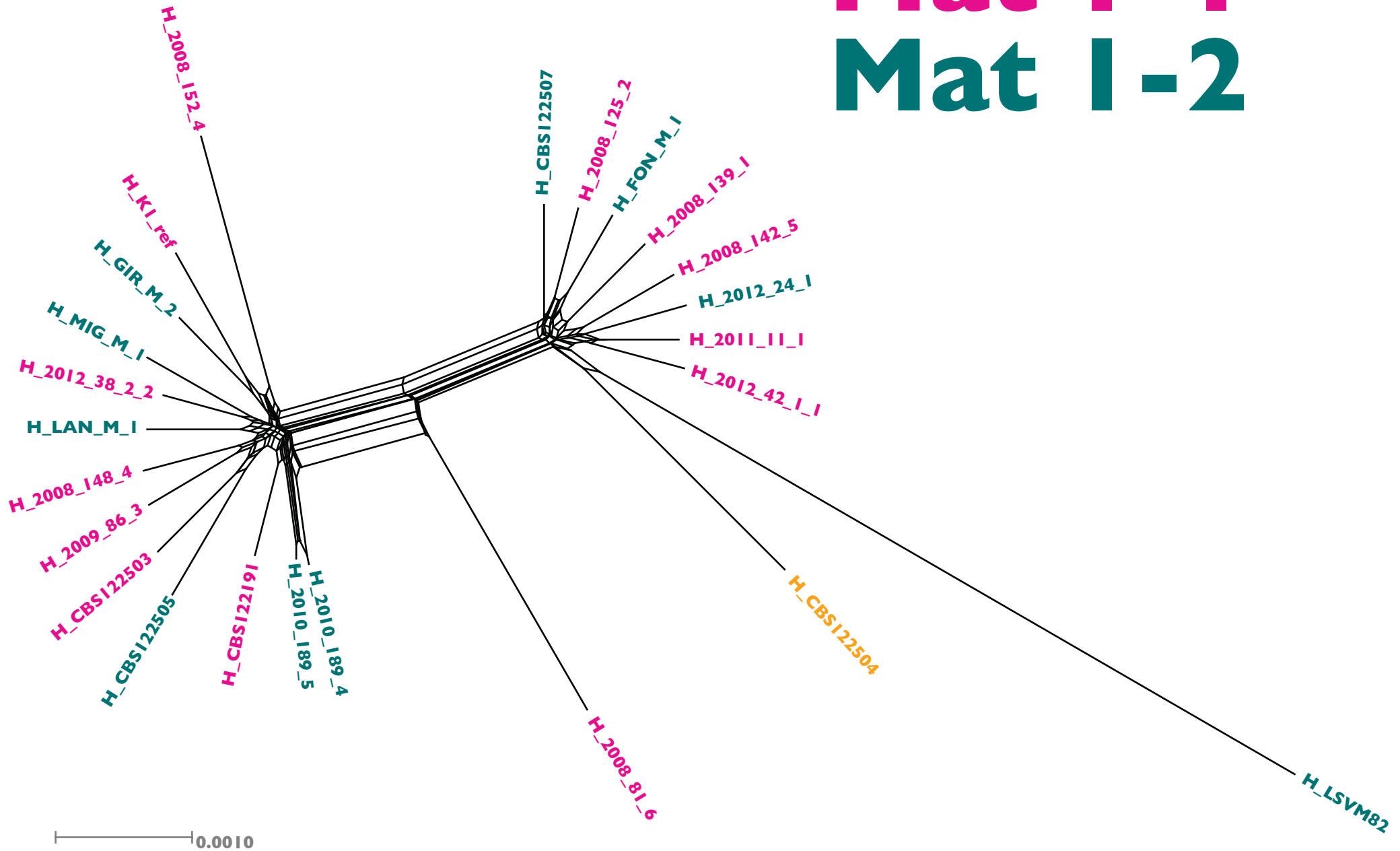


H. pseudoalbidus from Europe

NORNEX ID	MAT type	COUNTRY
H_2008_125_2	MAT-I-I	NORWAY
H_2008_139_1	MAT-I-I	NORWAY
H_2008_142_5	MAT-I-I	NORWAY
H_2008_148_4	MAT-I-I	NORWAY
H_2008_152_4	MAT-I-I	NORWAY
H_2008_81_6	MAT-I-I	NORWAY
H_2009_86_3	MAT-I-I	NORWAY
H_2010_189_4	MAT-I-2	NORWAY
H_2010_189_5	MAT-I-2	NORWAY
H_2011_11_1	MAT-I-I	NORWAY
H_2012_24_1	MAT-I-I	NORWAY
H_2012_38_2_2	MAT-I-I	NORWAY
H_2012_42_1_1	MAT-I-I	NORWAY
H_CBS122191	MAT-I-I	AUSTRIA
H_CBS122503	MAT-I-I	POLAND
H_CBS122504	nd	POLAND
H_CBS122505	MAT-I-2	POLAND
H_CBS122507	MAT-I-2	POLAND
H_FON_M_1	MAT-I-2	FRANCE
H_GIR_M_2	MAT-I-2	FRANCE
H_LAN_M_1	MAT-I-2	FRANCE
H_MIG_M_1	MAT-I-2	FRANCE
H_L SVM82	MAT-I-2	FRANCE

Mat I-1

Mat I-2



Revised aims

- What is the diversity of *H. pseudoalbidus* in Europe?
- Was there one introduction or many?
- Are strains interbreeding?
- Are there patterns of diversity with time of isolation in European *H. pseudoalbidus*?
- Is there association of genomic diversity with geography (and thus perhaps ash host genotype)?
- What was/were the origin(s) of the UK infections?

Questions, analyses, questions (I)

- 1 Validate assemblies, in particular screening out any "contaminants"
- 2 Validate variant calls
- 3 Repeat splitstree analyses and STRUCTURE to affirm initial findings
- 4 Investigate coalescent estimates of divergence and population sizes using genome-wide data
- 5 Integrate UK genome sequences into analysis pipeline
- 6 Explore patterns of divergence and conservation of specific genes and gene families

Questions, analyses, questions (II)

&& Acquire additional European samples for analyses exploring time and space in more detail

&& Produce reference quality assembly of "type" strain

NORNEX EDINBURGH

Georgios Koutsovoulos

