Listen to the title!

T. Ingalls, M. Marz, G. Martius, S. Prohaska

Leipzig

Bled, Slovenia
21.02.2009

History

- Put two writers of diploma thesis and dissertation in one room

History

- Put two writers of diploma thesis and dissertation in one room
- Let years go into the land, meet people

History

- Put two writers of diploma thesis and dissertation in one room
- Let years go into the land, meet people
- Write some very nice paper about it

First attempts

- tonics, subdominant, dominant, dominant seventh chord for censensus *CGAT*
  (programmed in `perl` by `MIDI::Music`)

First attempts

- tonics, subdominant, dominant, dominant seventh chord for censensus *CGAT*
  (programmed in `perl` by `MIDI::Music`)
  → cool

First attempts

- tonics, subdominant, dominant, dominant seventh chord for censensus *CGAT*
  (programmed in `perl` by `MIDI::Music`)
  - → cool
  - → boring

First attempts

- tonics, subdominant, dominant, dominant seventh chord for censensus *CGAT*
  (programmed in `perl` by `MIDI::Music`)
  - → cool
  - → boring
- without higher biological/artificial significance

Existing programs

	gene alignment	chromosomal alignment
sequence	`gene2music`, `PROMUSE`	`DNAmusic`
alignment	???	???

- `gene2music` automated conversion of protein-coding sequences (20 amino acids on 13 chords), grouping chemically similar characters together
- `PROMUSE` sonification of amino acid features with structural information and
- `DNAmusic` on large scale, less information transformed

Goal

- First tool: aligned chromosomal data $\longrightarrow$ musical composition

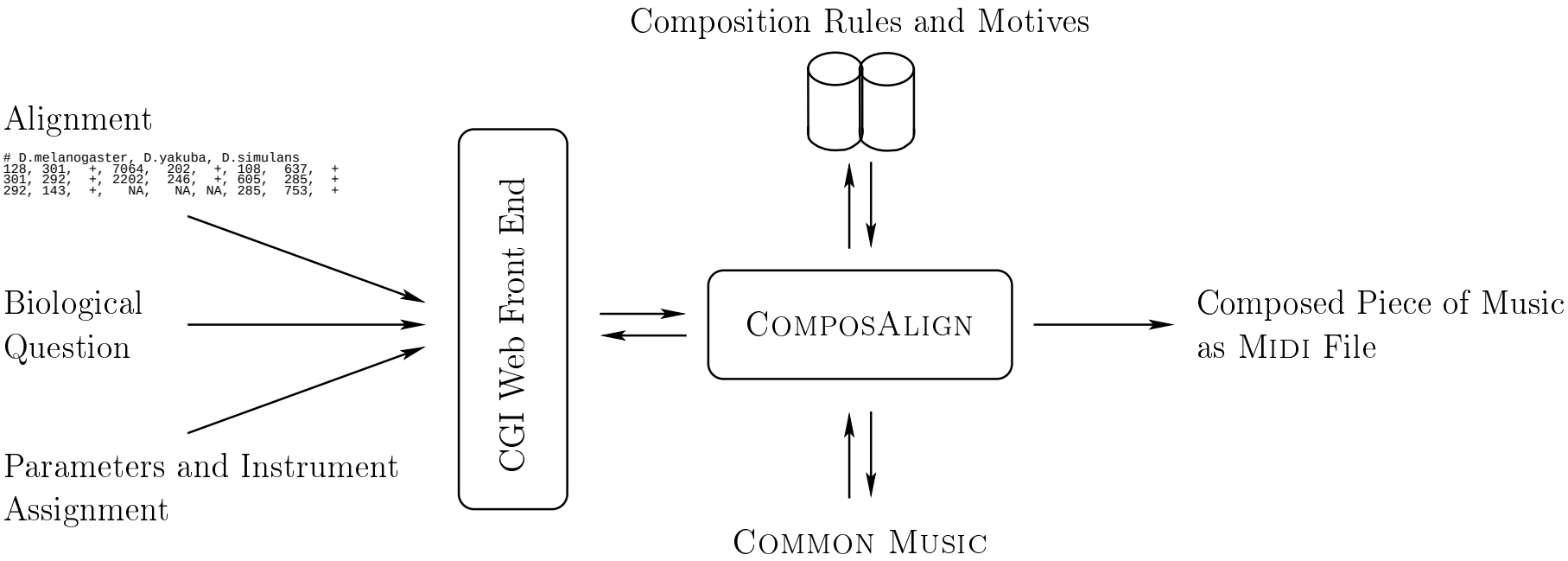
Goal

- First tool: aligned chromosomal data $\longrightarrow$ musical composition
- Transports information (answers to biological questions)

Goal

- First tool: aligned chromosomal data $\longrightarrow$ musical composition
- Transports information (answers to biological questions)
- Easy to understand, artistic and pleasant

Workflow


Composition Rules and Motives
Alignment
D.melanogaster, D.yakuba, D.simulans
128, 301, +, 7064, 202, +, 108, 637, +
301, 292, +, 2202, 246, +, 605, 285, +
292, 143, +, NA, NA, NA, 285, 753, +
Biological Question
Parameters and Instrument Assignment
CGI Web Front End
ComposAlign
Composed Piece of Music as Midi File
Common Music

Alignment input

```
#D.melanogaster, D.yakuba, D.simulans
1204181,1208982,+,1592422,1597568,+,203497,218425,+
1208982,1220029,+,1597568,1611936,+,1227473,1238543,+
1220029,1223495,+,1611936,1615470,+,1238543,1251804,+
1223495,1226280,+,1615470,1618686,+,408732,409578,-
1226280,1291345,+,1618686,1699795,+,1263986,1325900,+
1291345,1293751,+,NA,NA,NA,NA,NA,NA
1293751,1300981,+,1699795,1707089,+,1325900,1332929,+
1300981,1311360,+,1707089,1718160,+,1332929,1343270,+
1311360,1336219,+,1718160,1744017,+,1343270,1369930,+
1336219,1351937,+,1744017,1762657,+,1369930,1388290,+
```

Biological questions

Biological questions

- number of individuals
- occurence/absense of genes in a taxon
- occurence/absense of genes in certain subtrees
- direction of a genes
- distance of organisms
- distances of genes (repeats/clusters)

Biological questions

- number of individuals
- occurence/absense of genes in a taxon
- occurence/absense of genes in certain subtrees
- direction of a genes
- distance of organisms
- distances of genes (repeats/clusters)

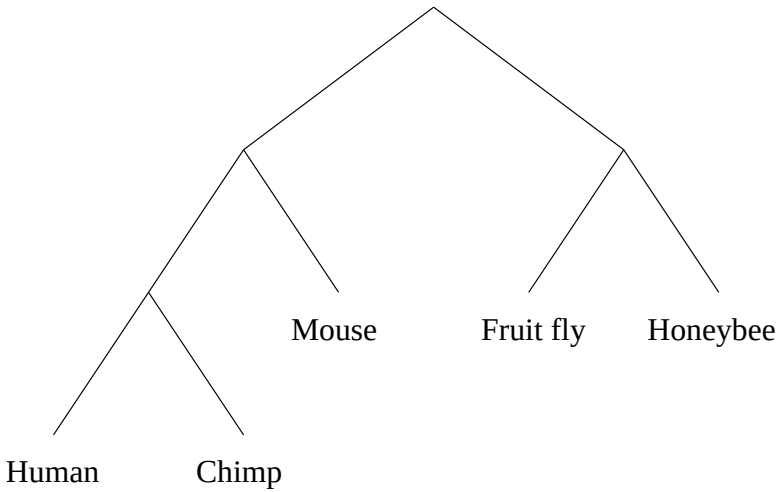
- TODO: local alignments code for structure information

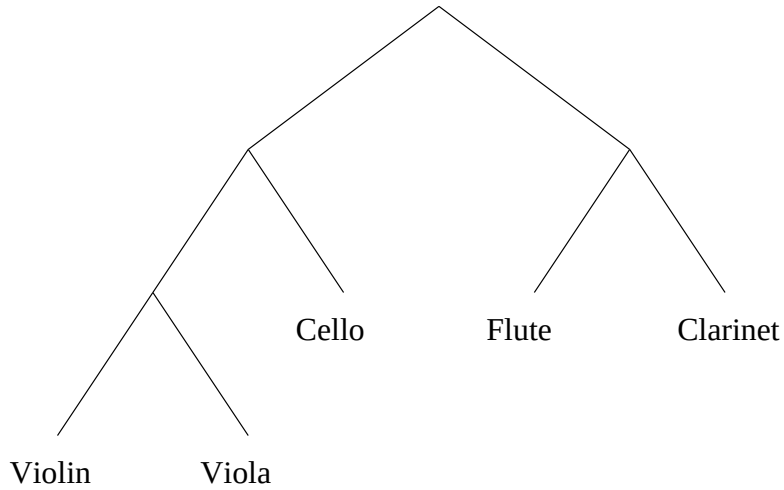
Mapping individuals

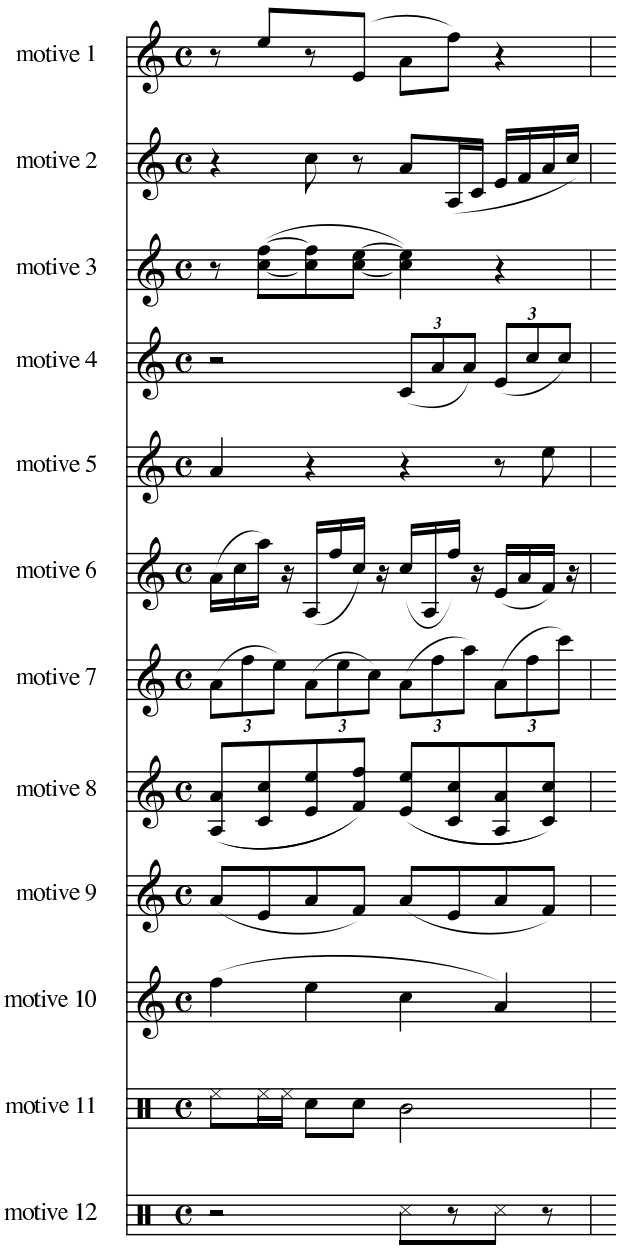
- $f : Ind \longrightarrow Ins$

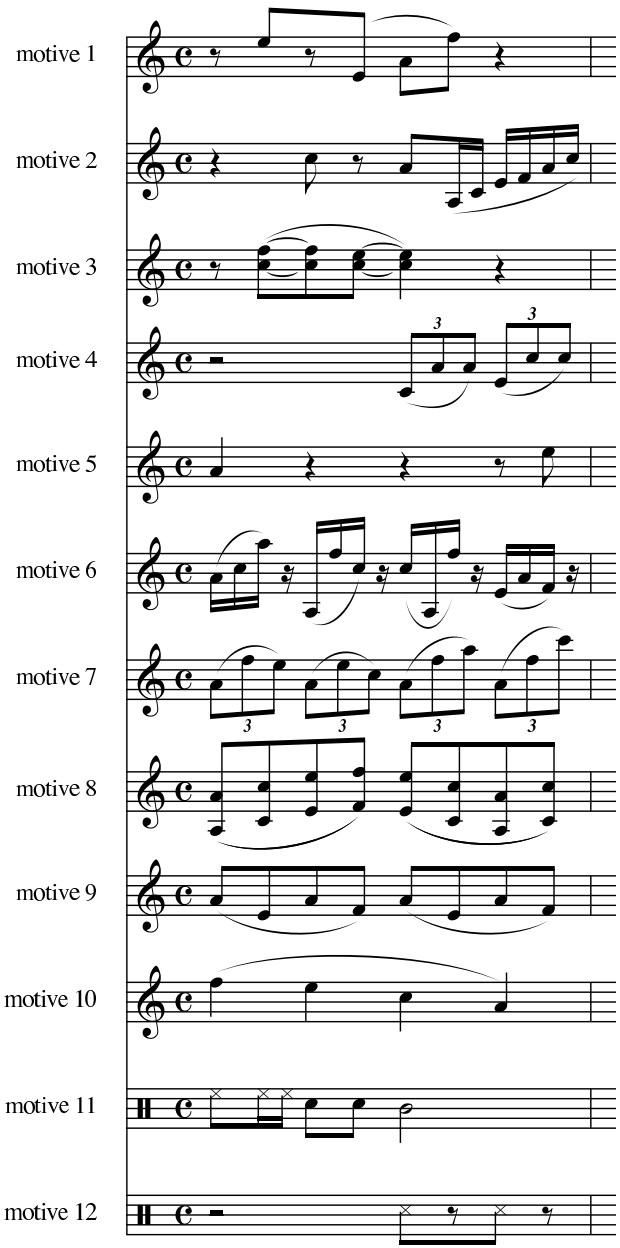
Mapping individuals

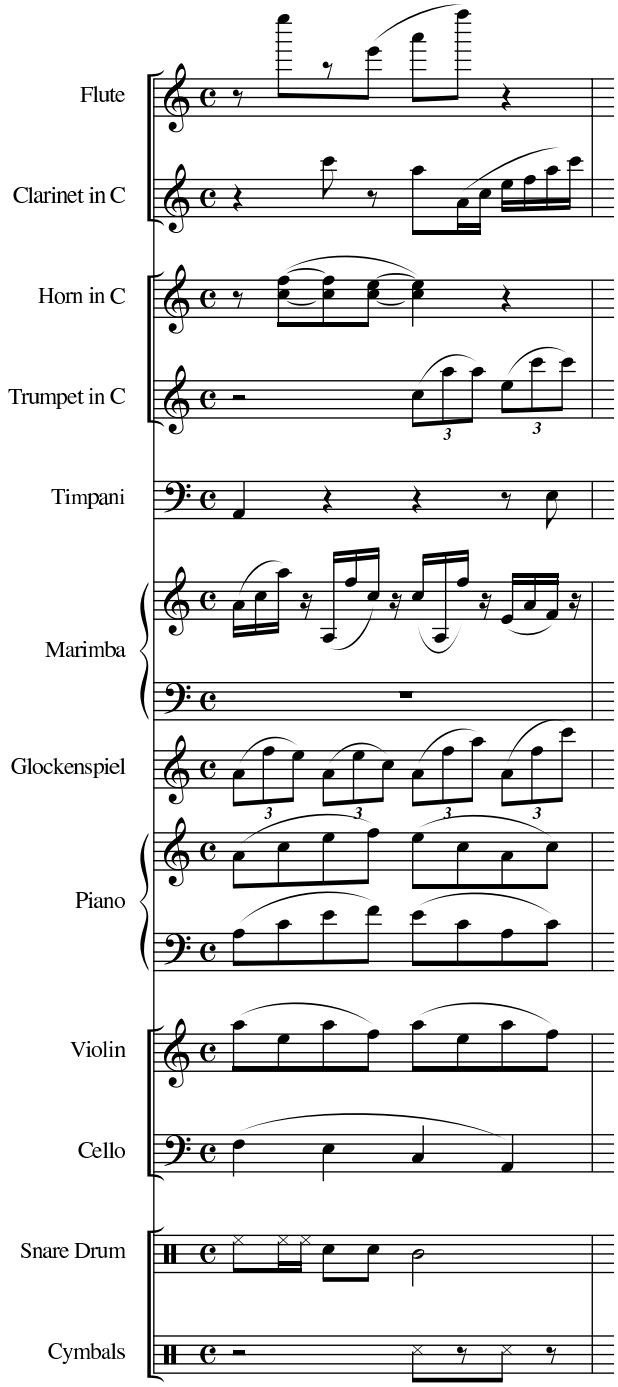
- $f : Ind \longrightarrow Ins$
- tree dependent

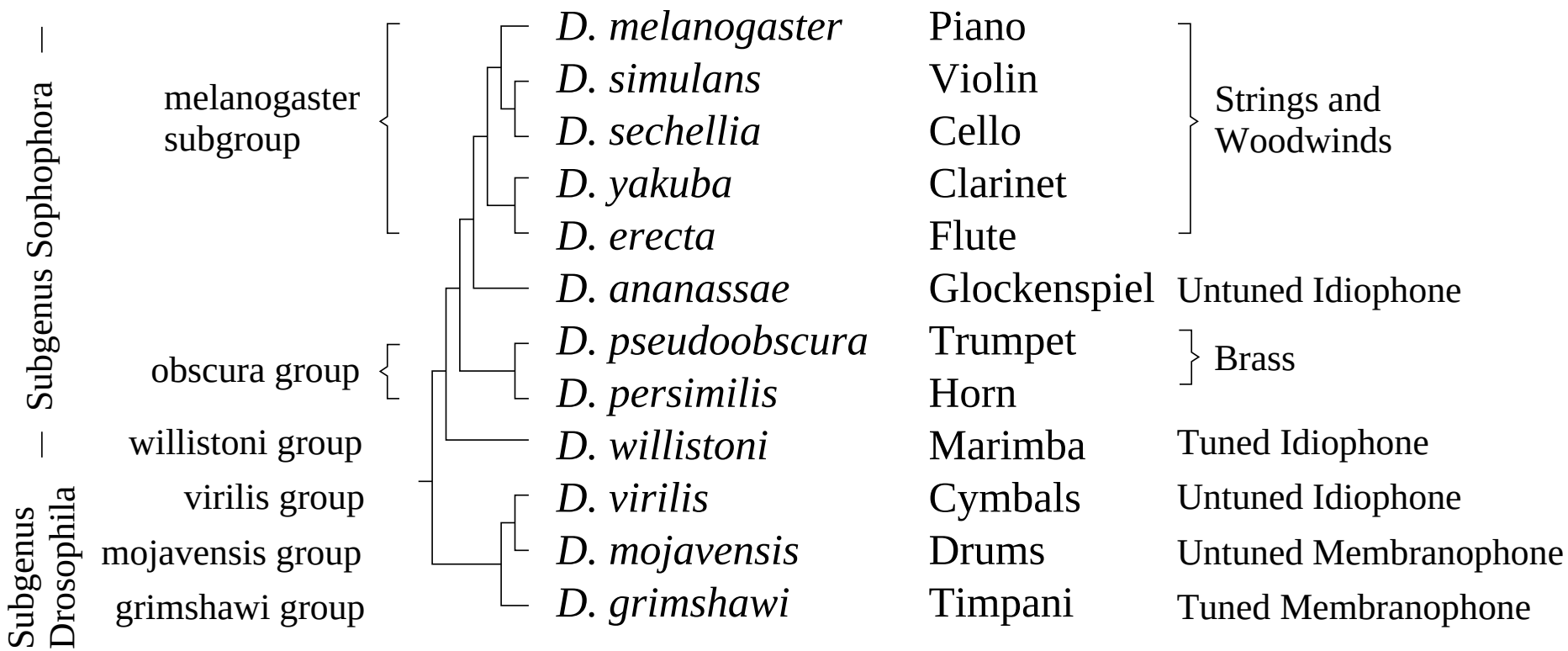





motive 1
motive 2
motive 3
motive 4
motive 5
motive 6
motive 7
motive 8
motive 9
motive 10
motive 11
motive 12

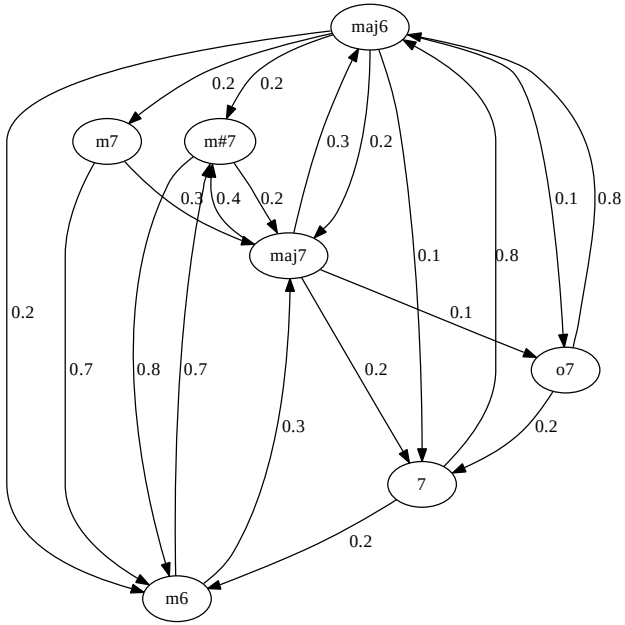

motive 1
motive 2
motive 3
motive 4
motive 5
motive 6
motive 7
motive 8
motive 9
motive 10
motive 11
motive 12


Flute
Clarinet in C
Horn in C
Trumpet in C
Timpani
Marimba
Glockenspiel
Piano
Violin
Cello
Snare Drum
Cymbals


Subgenus Drosophila
Subgenus Sophophora
melanogaster subgroup
obscura group
willistoni group
virilis group
mojavensis group
grimshawi group
D. melanogaster
D. simulans
D. sechellia
D. yakuba
D. erecta
D. ananassae
D. pseudoobscura
D. persimilis
D. willistoni
D. virilis
D. mojavensis
D. grimshawi
Piano
Violin
Cello
Clarinet
Flute
Glockenspiel
Trumpet
Horn
Marimba
Cymbals
Drums
Timpani
Strings and Woodwinds
Untuned Idiophone
Brass
Tuned Idiophone
Untuned Idiophone
Untuned Membranophone
Tuned Membranophone

How to improve

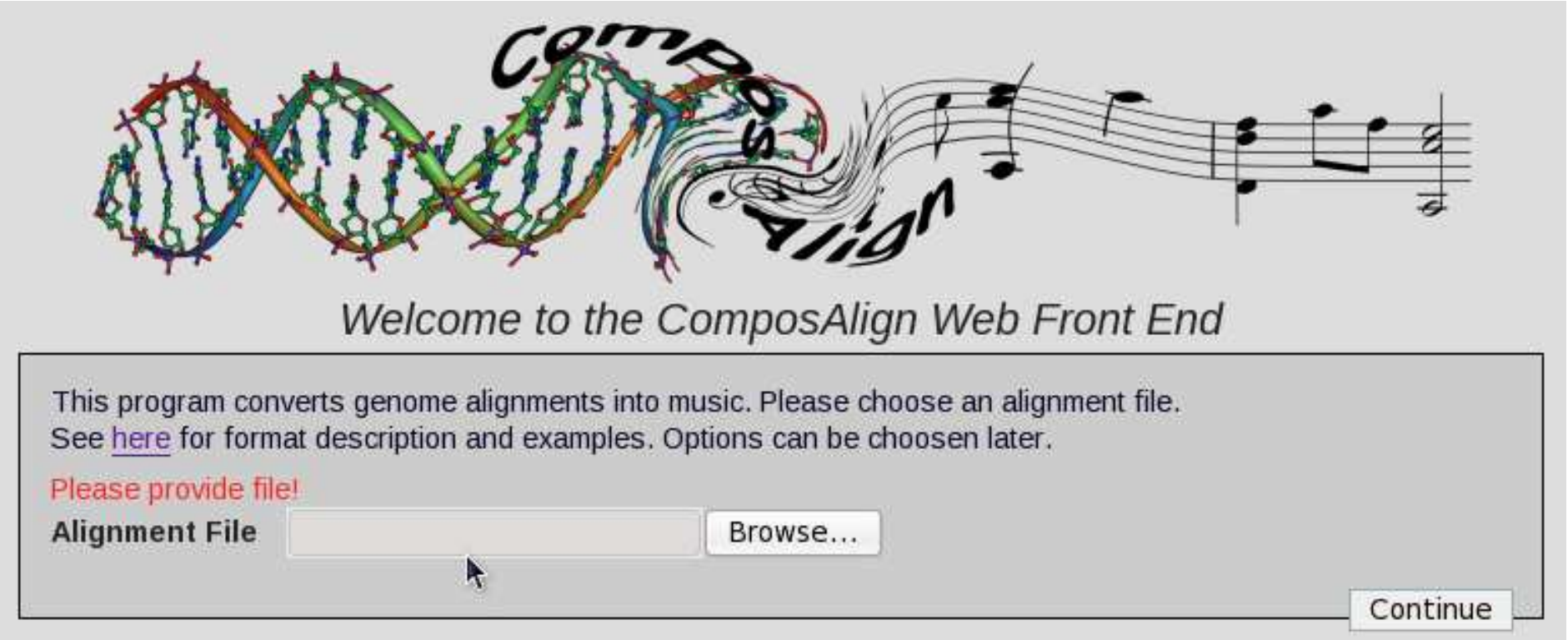
- maj6 – Major tonics sixth
- m#7 – Minor tonics seventh
- m7 – Minor diminished seventh
- maj7 – Major tonics seventh
- 7 – Major dominant
- m6 – Minor tonics sixth
- o7 – Minor dominant seventh



	maj6	*m#7*	*m7*	*maj7*	7	*m6*	*o7*
maj6	0	0.2	0.2	0.2	0.1	0.2	0.1
m#7	0	0	0	0.2	0	0.8	0
m7	0	0	0	0.3	0	0.7	0
maj7	0.3	0.4	0	0	0.2	0	0.1
7	0.8	0	0	0	0	0.2	0
m6	0	0.7	0	0.3	0	0	0
o7	0.8	0	0	0	0.2	0	0

Web Interface

http://www2.bioinf.uni-leipzig.de/cgi-bin/ComposAlign/ComposAlign.cgi



Goto interactive web front end here!

Example input-files for global alignment:

- Chromosome 3R of 3 flies
- Chromosome 3R of 12 flies

Example output-files for global alignment:

Example files with three flies:

- Orientation not translated, generated without a markov model
  [mid] [mp3]
- Orientation not translated, generated with a markov model
  [mid] [mp3]
- Orientation translated, generated with a markov model
  [mid] [mp3]

12 flies, all examples genereated with markov model

- Orientation not translated
  [mid] [mp3]
- Orientation not translated, compressed accords for completely aligned regions
  [mid] [mp3]
- Orientation translated, compressed accords for completely aligned regions
  composed with different mappings
  [mid] [mp3]
  [mid] [mp3]
  [mid] [mp3]
  [mid] [mp3]
  [mid] [mp3]
  [mid] [mp3]

Format Description of input files:
A first line starting with '#' is the organisms description line.
Other lines present an alignment for a single gene.
Lines are ',' separated.
Three coloumns for each organism:
STARTPOSITION OF A GENE, STOPPOSITION OF A GENE, ORIENTATION OF A GENE

Alignment File `all.R3.dir.map`

Biological Question: global alignment

I detected 12 species in your file.

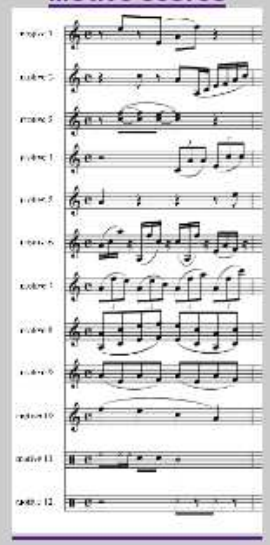
You have now the possibility to change some parameters for the music generation. It is possible to disable certain species by setting the instrument to `Disabled`.

Reference Species? 1

Sort data? ☑

Species	Instruments	Patterns
1	piano	motive8
2	violin	motive9
3	cello	motive10
4	glockenspiel	motive7
5	flute	motive1
6	timpani	motive5
7	drums	drums
8	horn	motive3
9	trumpet	motive4
10	clarinet	motive2
11	cymbal	cymbal
12	marimba	motive6

motive scores



Strand important? ☑

Conserved parts as accords? ☑

Harmonic changes? ☑

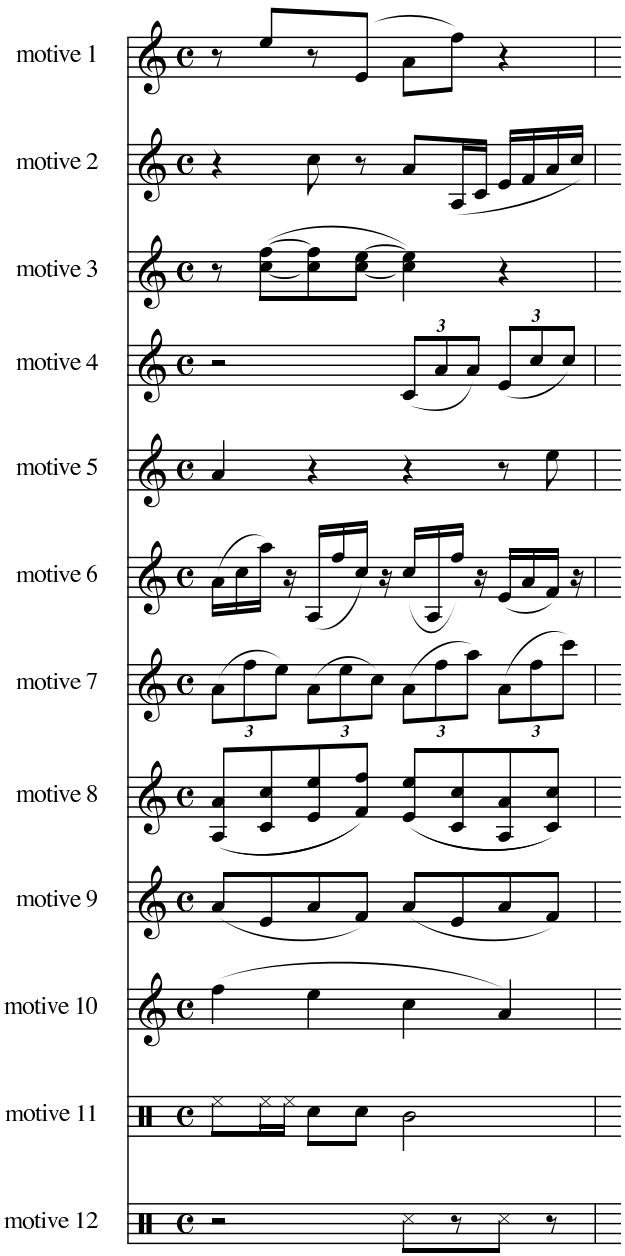
Generate Music

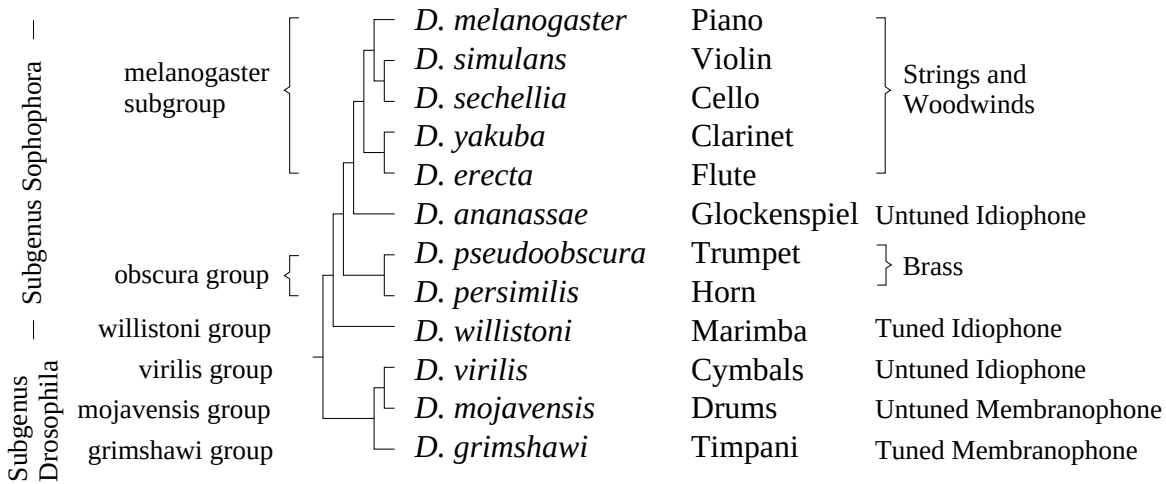
Acknowledgements

Thx 2:

Todd Ingalls, Georg Martius, Sonja Prohaska
Peter Stadler
Petra Pregel, Jens Steuck
and the whole bioinformatics group leipzig

Thank You!


motive 1
motive 2
motive 3
motive 4
motive 5
motive 6
motive 7
motive 8
motive 9
motive 10
motive 11
motive 12


Subgenus Drosophila
Subgenus Sophophora
melanogaster subgroup
obscura group
willistoni group
virilis group
mojavensis group
grimshawi group
D. melanogaster
D. simulans
D. sechellia
D. yakuba
D. erecta
D. ananassae
D. pseudoobscura
D. persimilis
D. willistoni
D. virilis
D. mojavensis
D. grimshawi
Piano
Violin
Cello
Clarinet
Flute
Glockenspiel
Trumpet
Horn
Marimba
Cymbals
Drums
Timpani
Strings and Woodwinds
Untuned Idiophone
Brass
Tuned Idiophone
Untuned Idiophone
Untuned Membranophone
Tuned Membranophone