

ВІОНАСК 2018

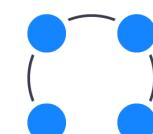
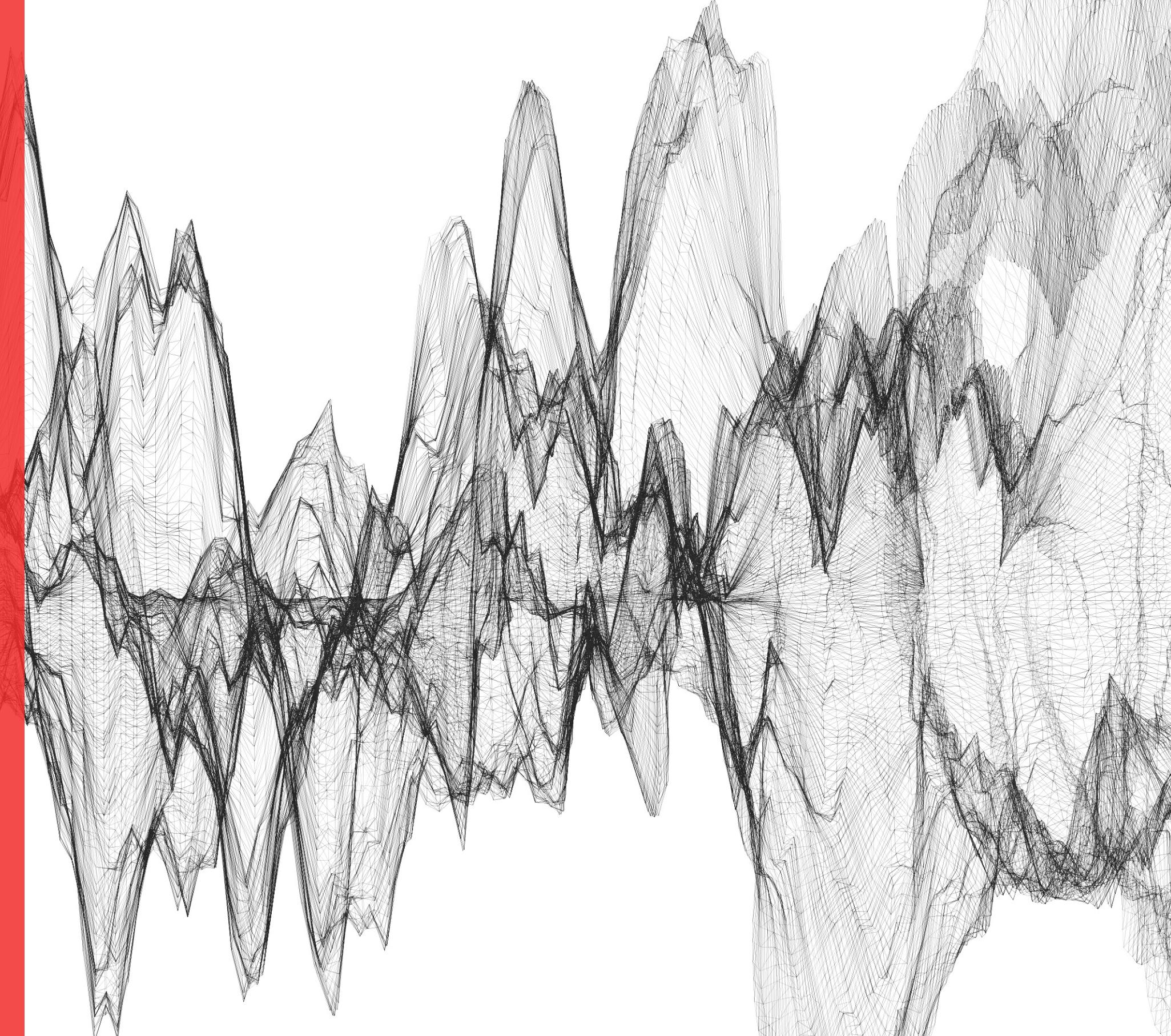
АНАЛИЗ ЗВУКОПРОДУКЦІИ ДРОЗОФИЛЫ

GOOD VIBRATIONS

- Наумов Станислав
- Кислюк Александр
- Рабдано Севастьян
- Беломестных Сергей
- Быков Кирилл

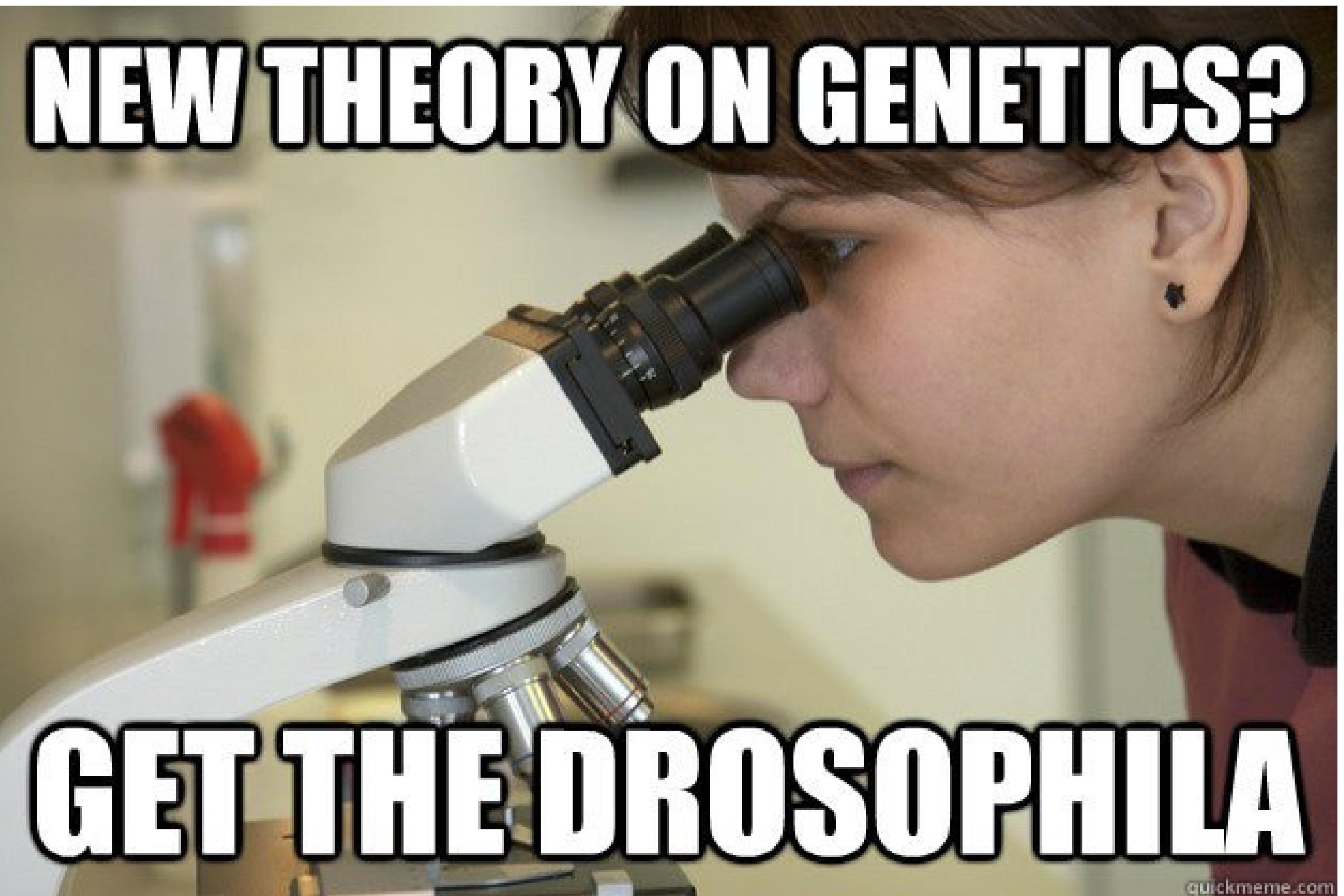
ЭКСПЕРТ:

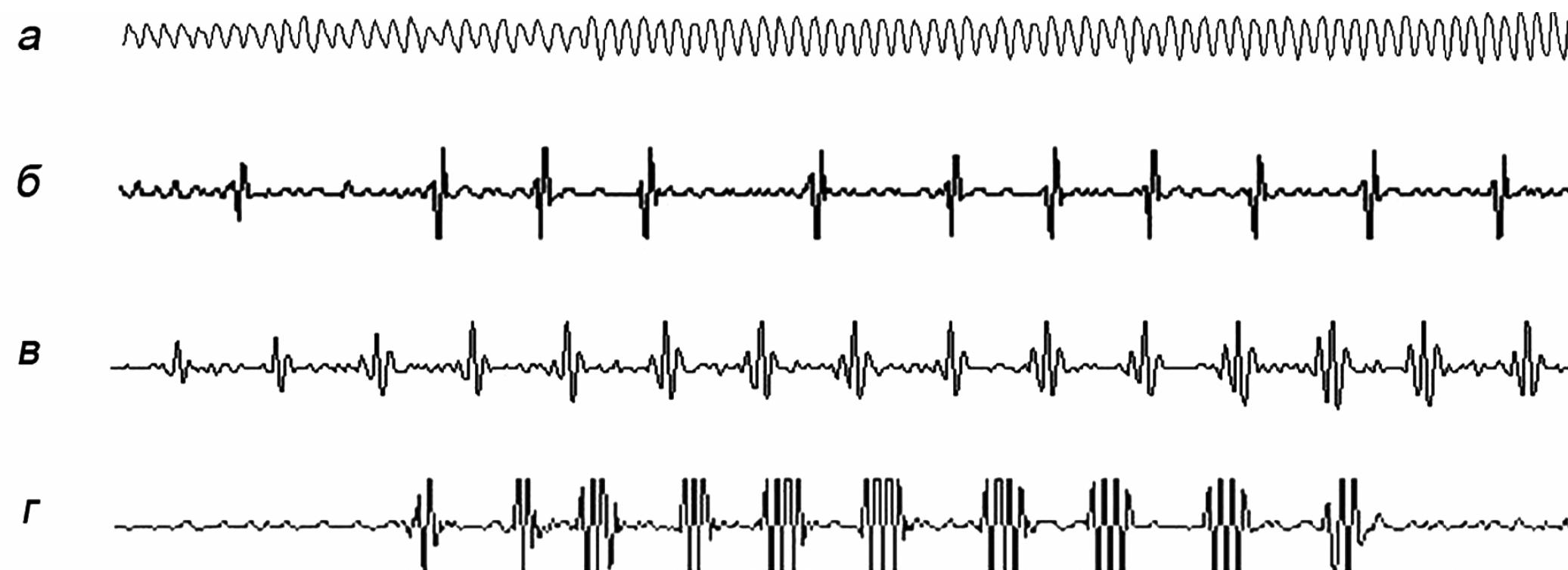
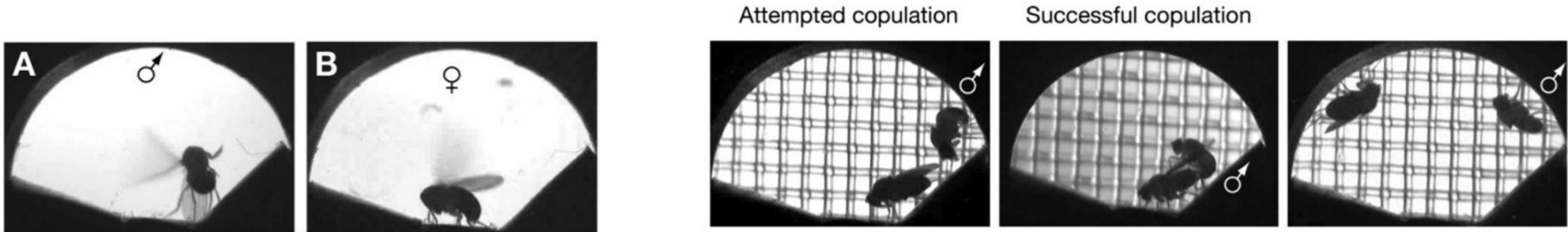
- Геннадий Захаров



NEW THEORY ON GENETICS?

GET THE DROSOPHILA



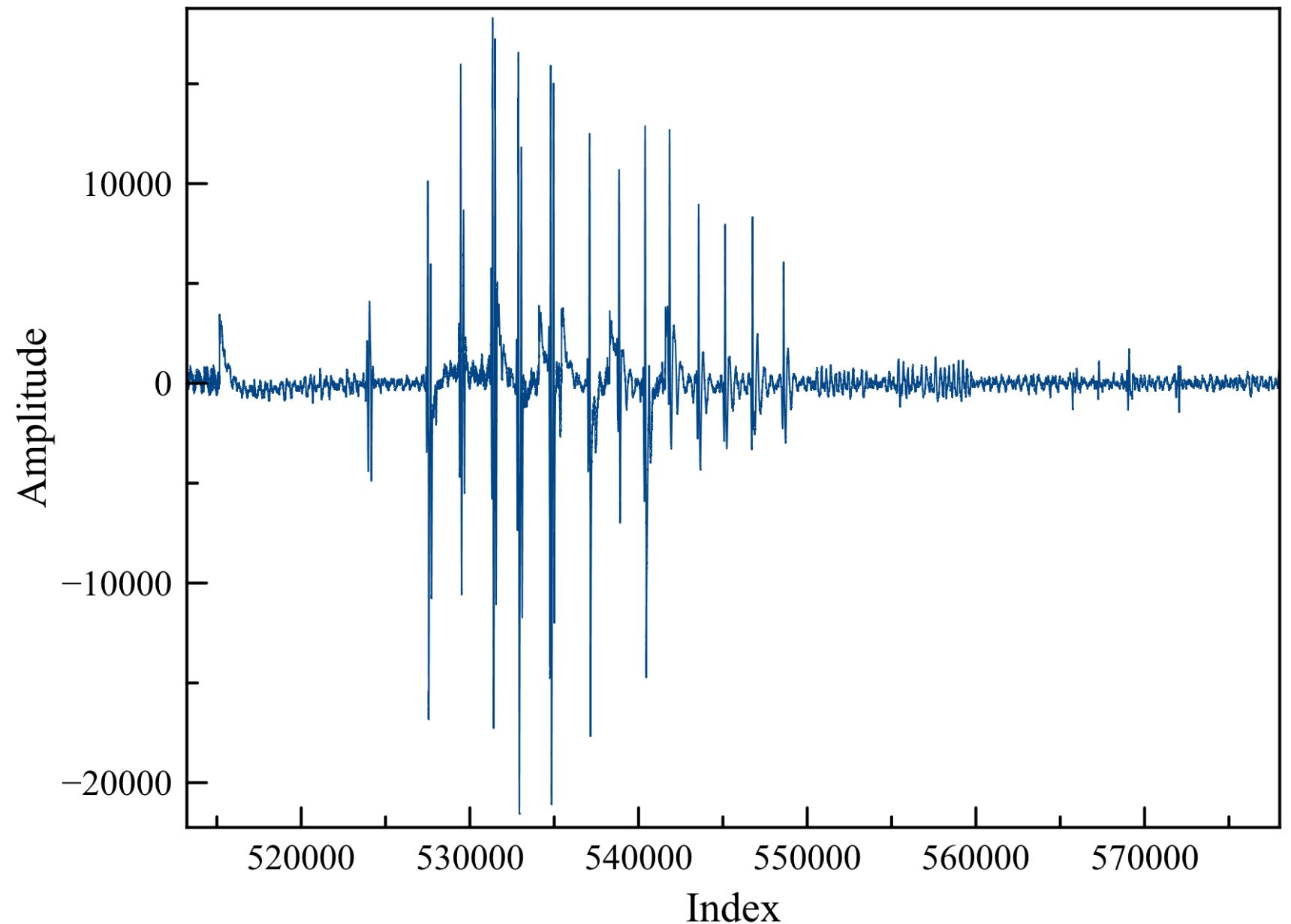


Осциллограммы сигналов ухаживания самцов *D.melanogaster* синусоидальной песни (а) и импульсной песни с нормальными (б) и искаженными пульсами (в, г)

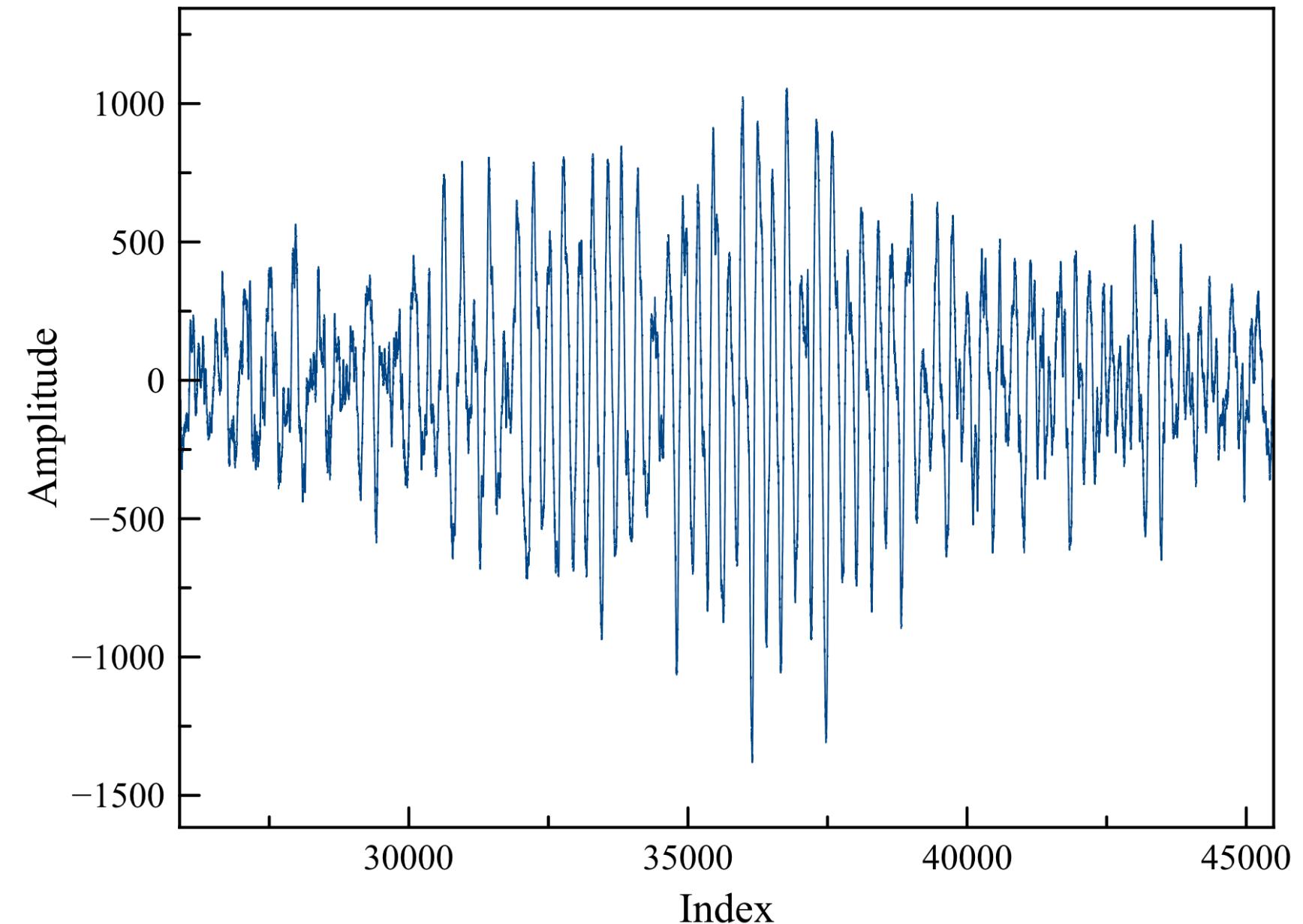
ЗАДАЧИ

- Создание утилиты для автоматической разметки и аннотации аудиофайлов
- Классификация импульсных и синусоидальных сигналов
- Извлечение параметров песни
- Поиск статистически значимых различий в параметрах песен для особей дикого типа и мутантов

Импульсная песня

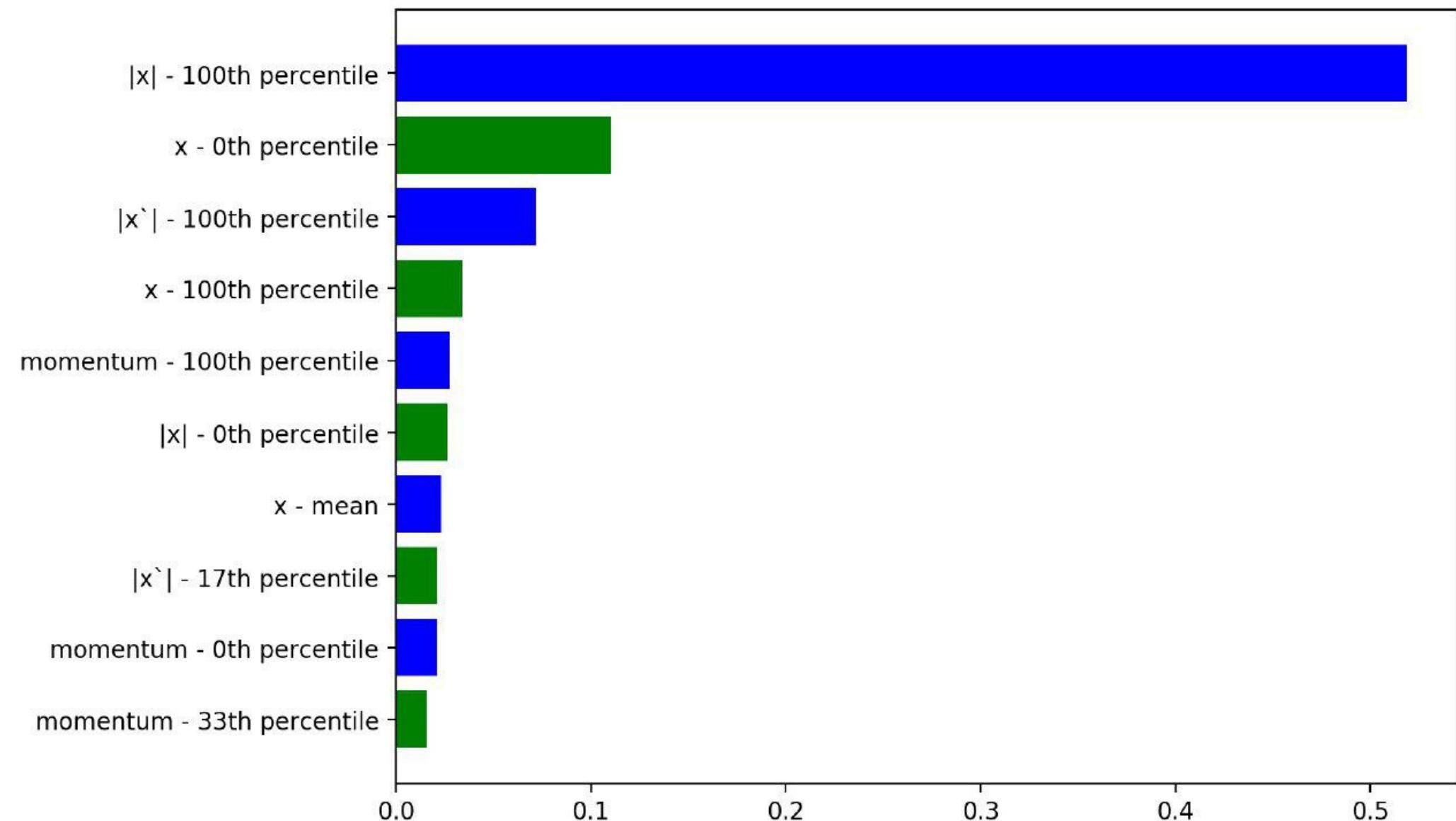


Синусоидальная песня



СЕГМЕНТАЦИЯ

- Разбиваем данные на окна, и вычисляем статистические признаки
- На основе аннотированных данных производится обучение модели градиентного бустинга
- При помощи обученной модели производится сегментация звуковой дорожки



plotly scikit-learn Dash

pandas $y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$

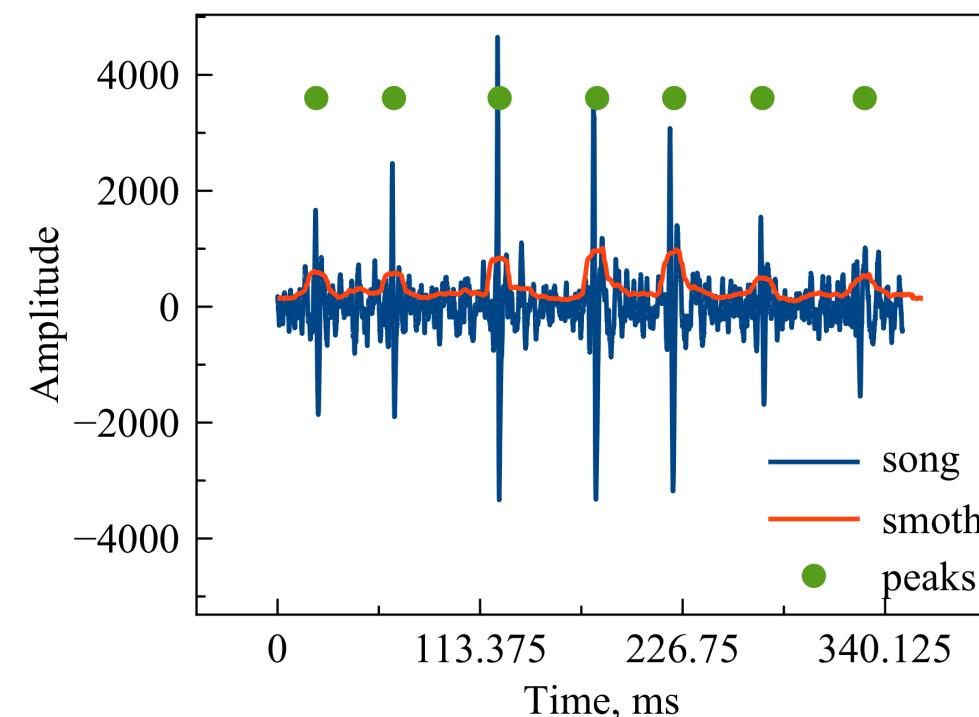
Значимость признаков

BIONACK 2018

ИЗВЛЕЧЕНИЕ ПАРАМЕТРОВ ПЕСЕН

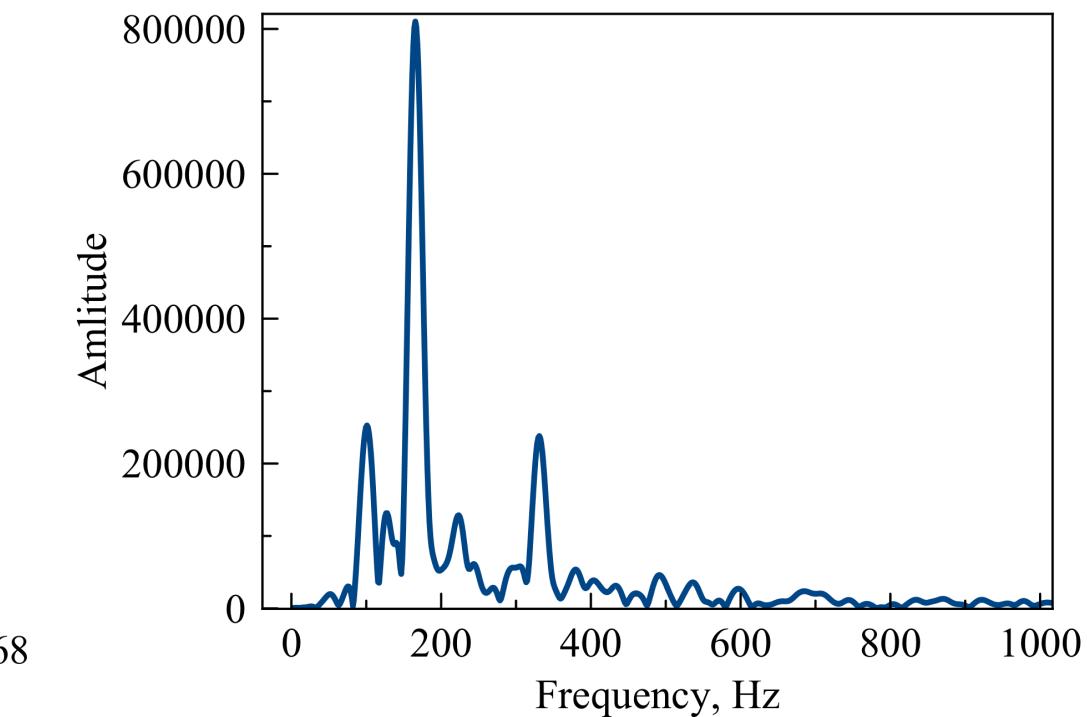
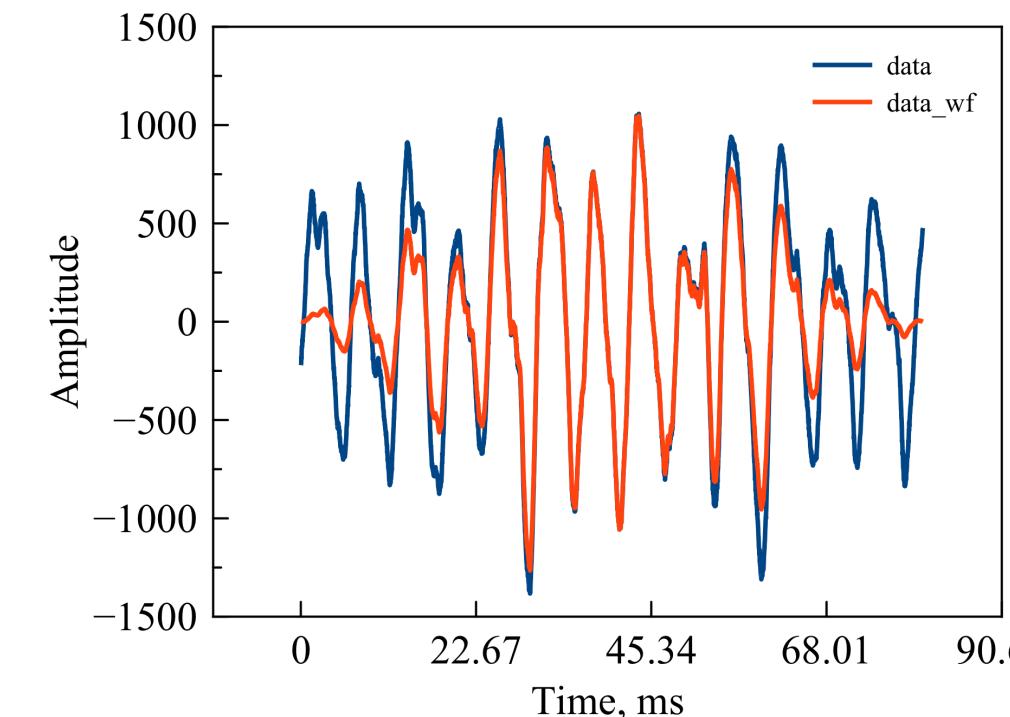
Импульсная песня

- Количество импульсов
- Длительность песни
- Частота следования импульсов
- Энергия импульса
- Максимальная амплитуда импульса
- Ширина импульса

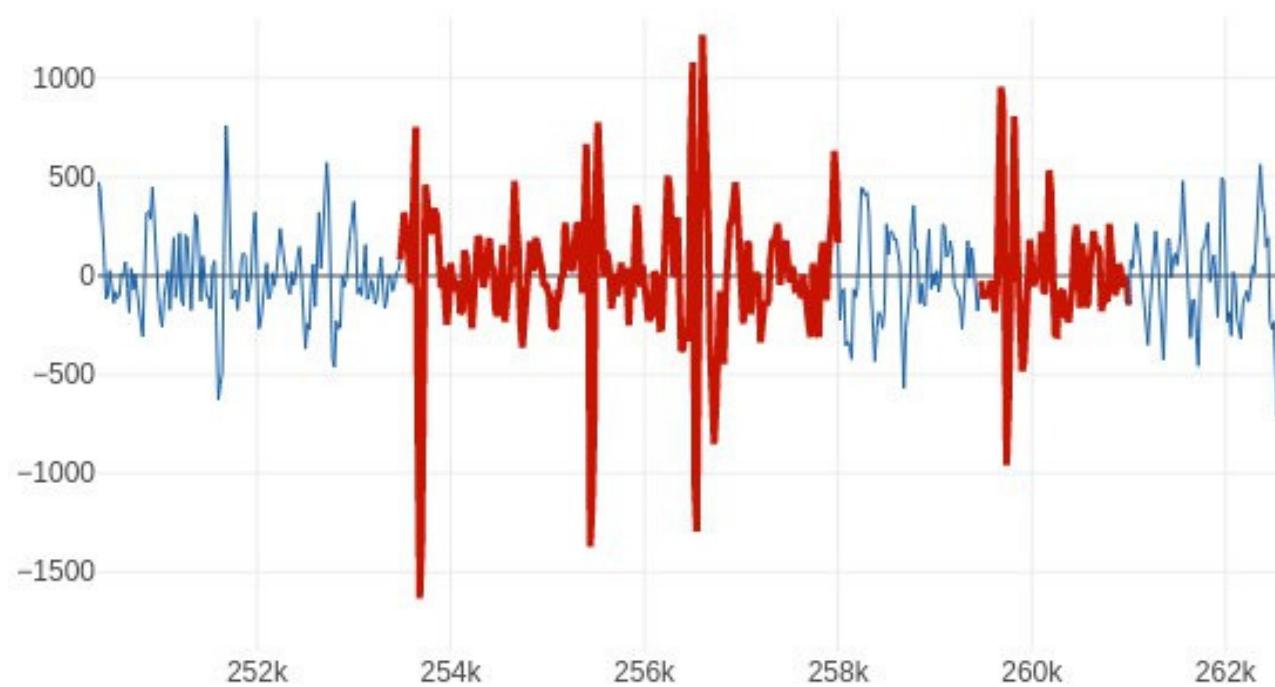


Синусоидальная песня

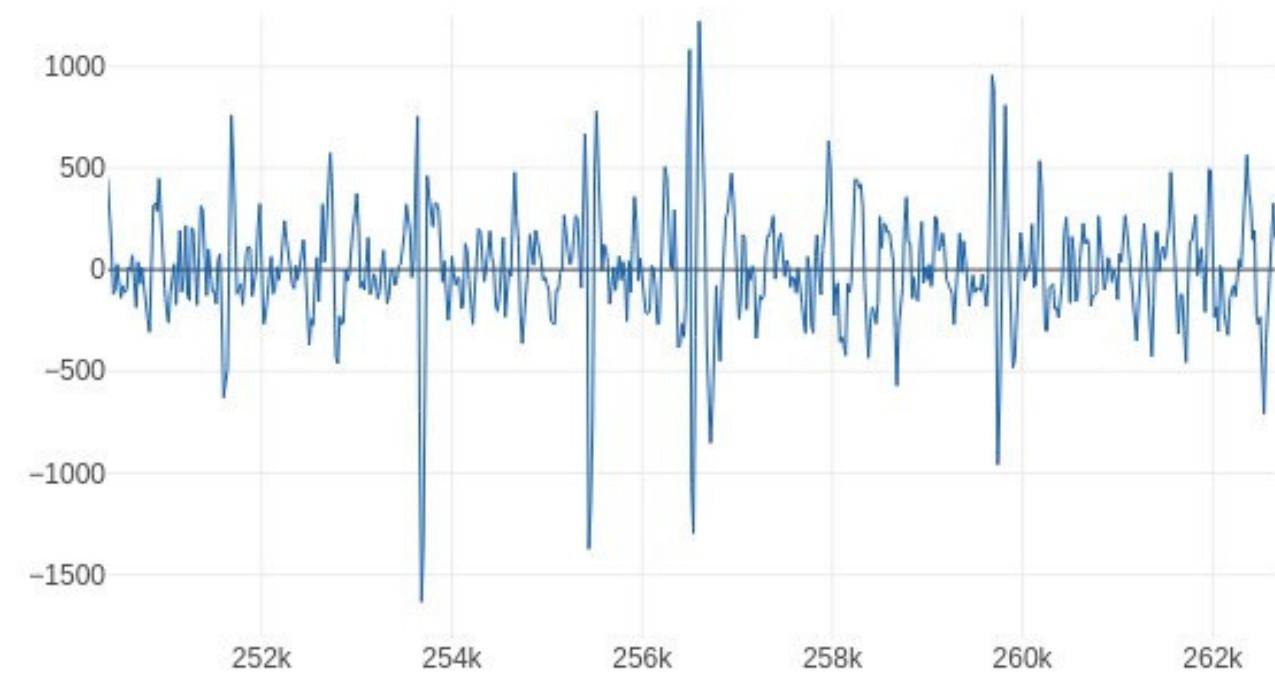
- Длительность песни
- Частота колебаний
- Количество периодов в песне
- Модуляция амплитуды



Сегментация нашим
алгоритмом



Сегментация человеком



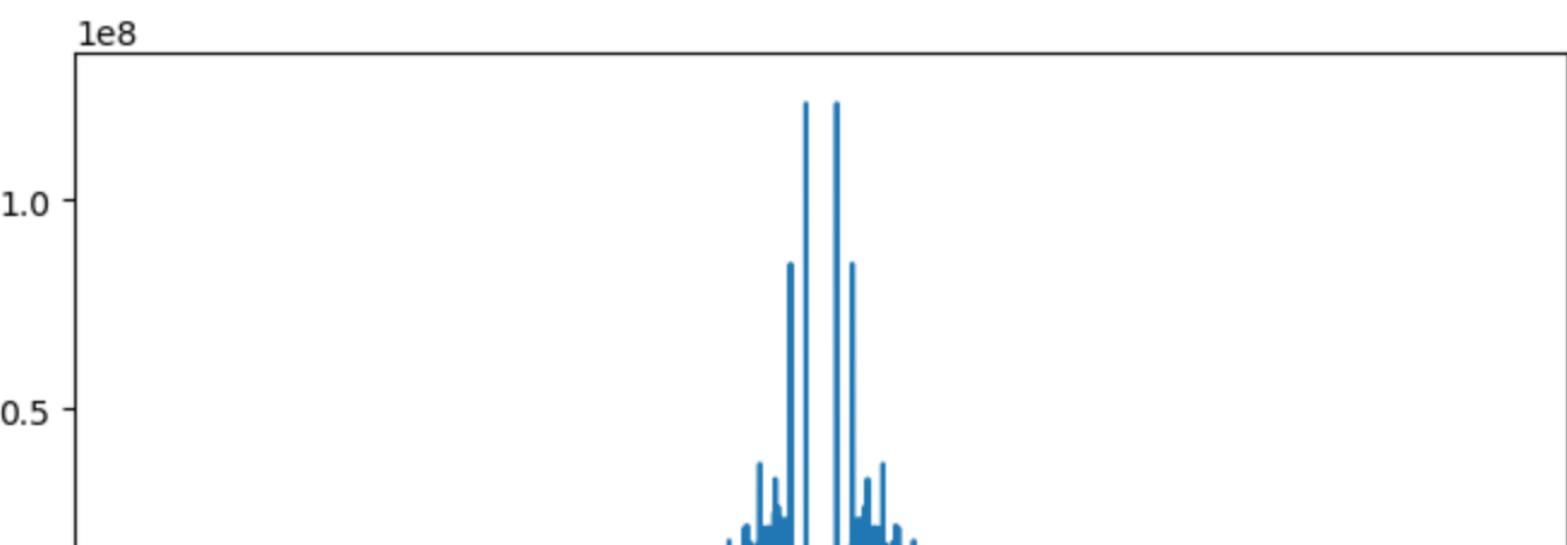
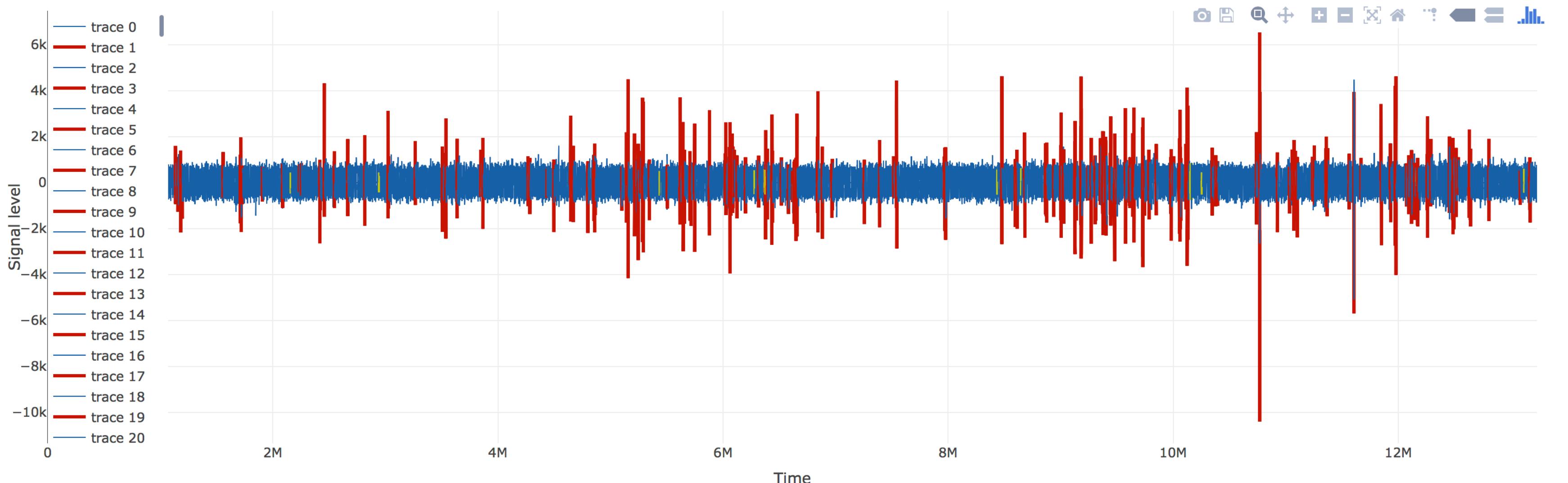


BioHack2018 — python < python utility.py -inp 22-03-16-agn-01.wav — 106x34

```
Last login: Sun Mar  4 10:07:42 on ttys000
[MacBook-Pro-Kirill:~ kirillbykov$ cd Documents/
[MacBook-Pro-Kirill:Documents kirillbykov$ cd GitHub/
[MacBook-Pro-Kirill:GitHub kirillbykov$ cd BioHack2018/
[MacBook-Pro-Kirill:BioHack2018 kirillbykov$ python utility.py -inp 22-03-16-agn-01.wav > logfile.txt
 * Running on http://127.0.0.1:8050/ (Press CTRL+C to quit)
 * Restarting with stat
 * Debugger is active!
 * Debugger PIN: 138-447-890
^C[MacBook-Pro-Kirill:BioHack2018 kirillbykov$ python utility.py -inp 22-03-16-ag01.wav > logfile.txt
 * Running on http://127.0.0.1:8050/ (Press CTRL+C to quit)
 * Restarting with stat
 * Debugger is active!
 * Debugger PIN: 138-447-890
127.0.0.1 - - [04/Mar/2018 10:19:28] "GET /_dash-layout HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:19:28] "GET /_dash-dependencies HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:19:28] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:19:28] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:20:06] "GET /_dash-layout HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:20:06] "GET /_dash-dependencies HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:20:06] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:22:12] "GET /_dash-layout HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:22:12] "GET /_dash-dependencies HTTP/1.1" 200 -
127.0.0.1 - - [04/Mar/2018 10:22:12] "GET / HTTP/1.1" 200 -
```

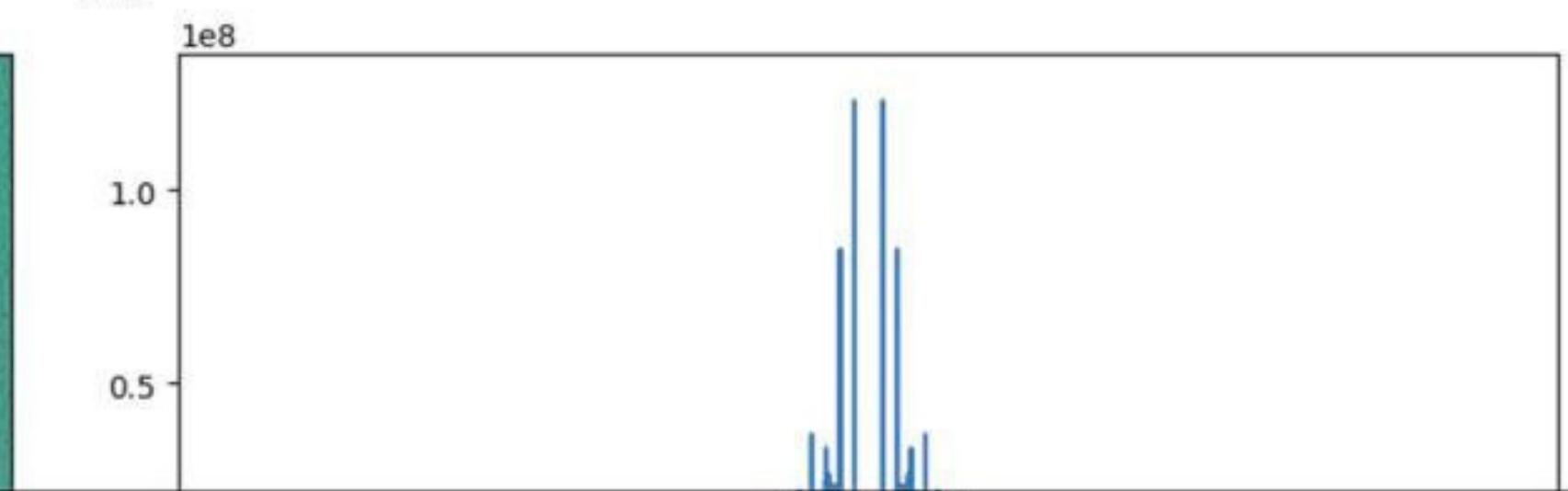
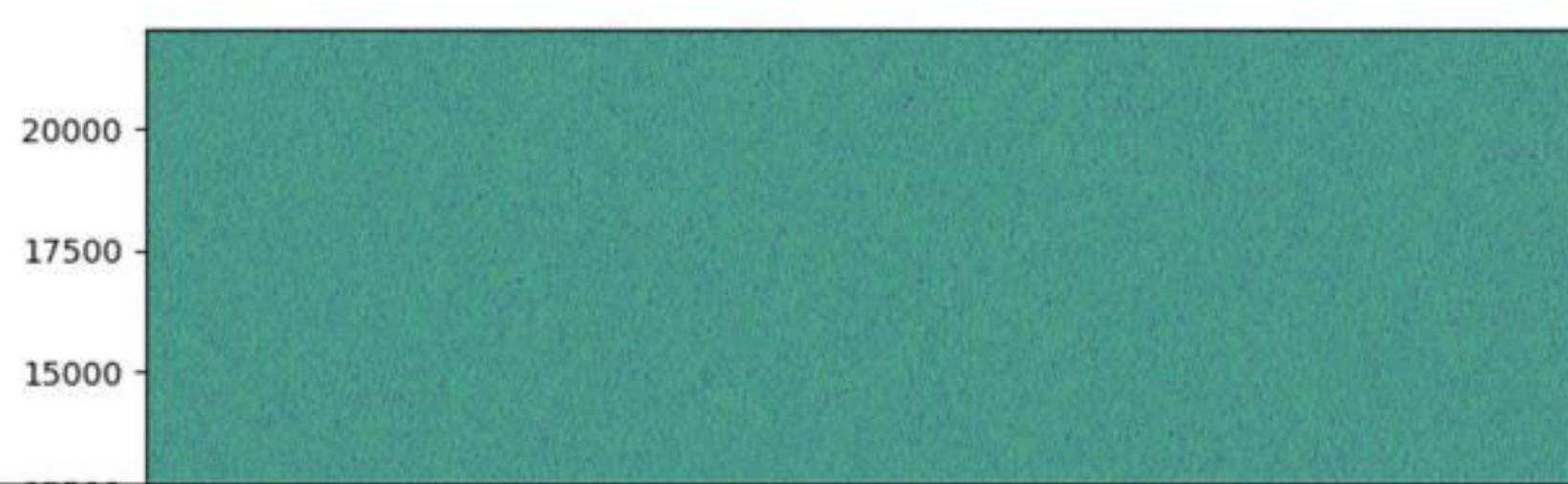
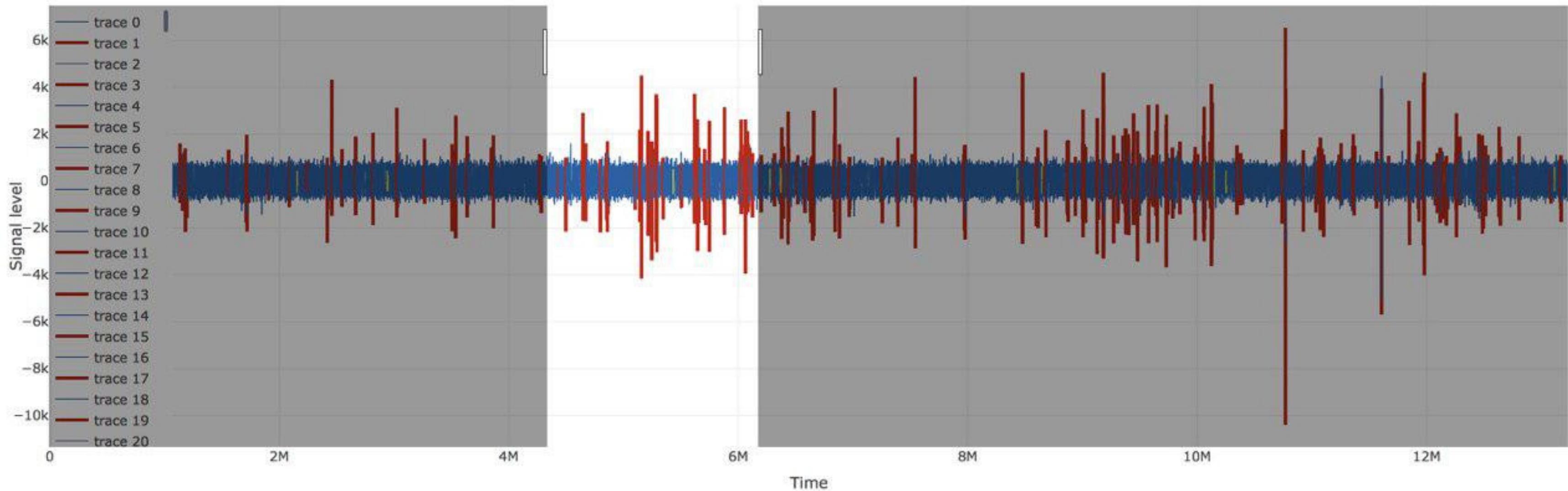
Analysis of Drosophila sound production

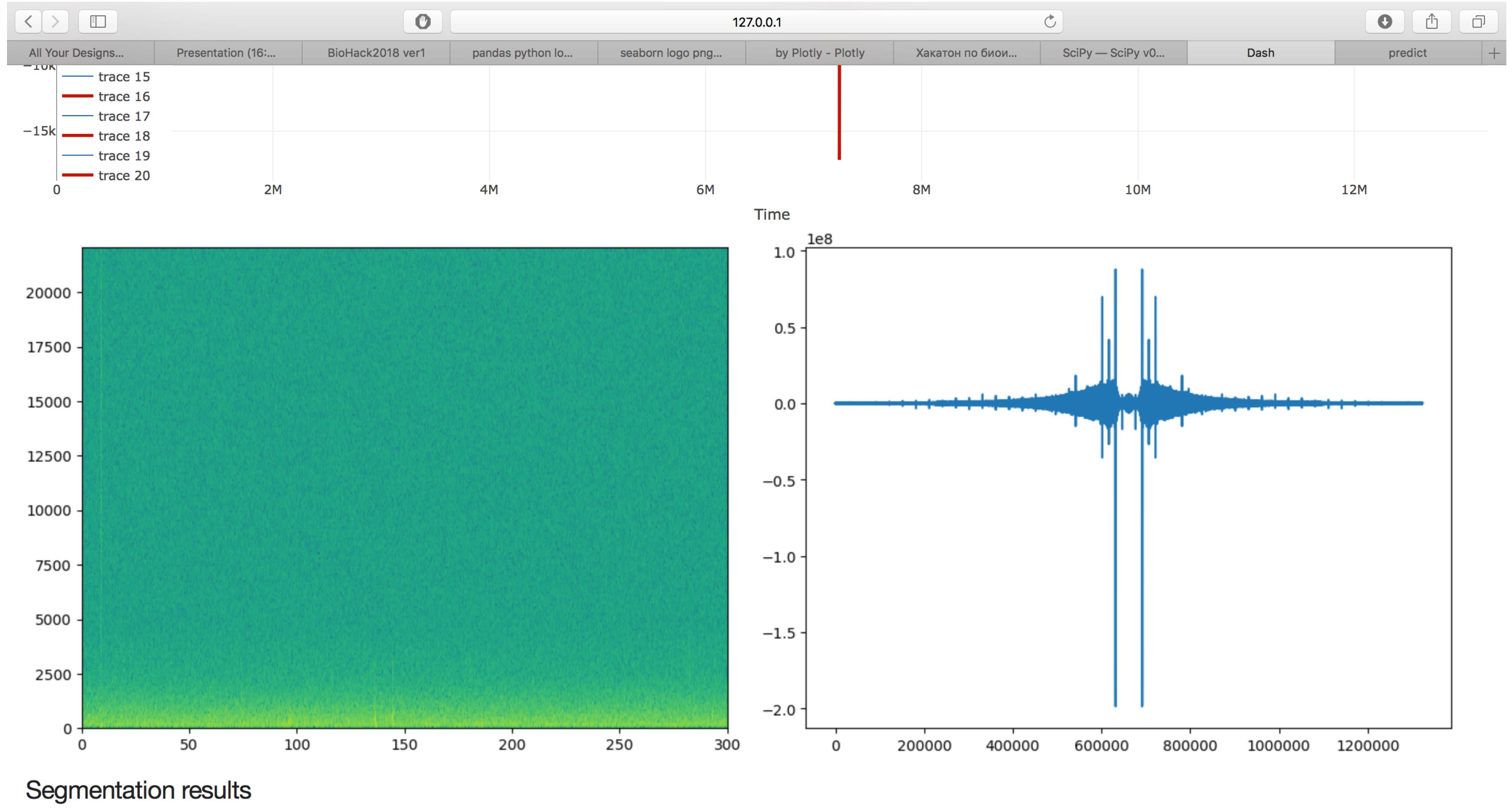
Good vibrations – BioHack 2018, Saint-Petersburg



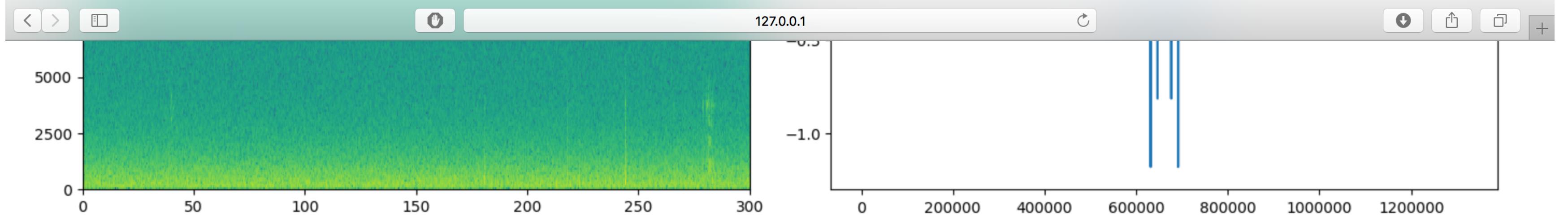
Analysis of Drosophila sound production

Good vibrations – BioHack 2018, Saint-Petersburg





FILTER ROWS



Segmentation results

FILTER ROWS

<input type="checkbox"/>	start	end	▲type
	<input type="text" value="Search"/>	<input type="text" value="Search"/>	<input type="text" value="s"/>
<input type="checkbox"/>	10158000	10159500	S
<input type="checkbox"/>	10155000	10156500	S
<input type="checkbox"/>	8653500	8655000	S
<input type="checkbox"/>	8440500	8442000	S
<input type="checkbox"/>	6381000	6382500	S
<input type="checkbox"/>	6373500	6375000	S
<input type="checkbox"/>	6283500	6285000	S
<input type="checkbox"/>	5440500	5442000	S
<input type="checkbox"/>	2949000	2950500	S
<input type="checkbox"/>	2161500	2163000	S

Pulse waves:

<input type="button" value="<"/>	<input type="button" value=">"/>	<input type="button" value="refresh"/>	127.0.0.1	<input type="button" value="refresh"/>	<input type="button" value="download"/>	<input type="button" value="upload"/>	<input type="button" value="new"/>	<input type="button" value="+"/>
2101000	2103000	3						

Pulse waves:

FILTER ROWS

<input type="checkbox"/>	start	end	number_of_pulses	song_duration	maxamps_mean	maxamps_std	widths_mean	widths_std	energies_mean	energies_std
<input type="checkbox"/>	1557000	1558500	1	0.0340136054421	1467	0	0.0256789453945	0	-21237	0
<input type="checkbox"/>	1710000	1719000	2	0.2040816326530	2128.5	92.5	0.0395270045505	0.0195270045505	6904	513
<input type="checkbox"/>	1728000	1729500	1	0.0340136054421	1234	0	0.0358740985517	0	-6674	0
<input type="checkbox"/>	1905000	1906500	1	0.0340136054421	1212	0	0.02	0	14863	0
<input type="checkbox"/>	2086500	2088000	1	0.0340136054421	1214	0	0.1388241197336	0	28676	0
<input type="checkbox"/>	2233500	2235000	1	0.0340136054421	1414	0	0.02	0	-11240	0
<input type="checkbox"/>	2253000	2254500	1	0.0340136054421	1241	0	0.0558510718009	0 0	3701	0
<input type="checkbox"/>	2415000	2421000	2	0.1360544217687	2068.5	604.5	0.0838163857244	0.0528150812553	1777.5	22795.5
<input type="checkbox"/>	2455500	2458500	1	0.0680272108843	4499	0	0.0548246631951	0	10440	0
<input type="checkbox"/>	2544000	2547000	1	0.0680272108843	1682	0	0.02	0	-20953	0
<input type="checkbox"/>	2664000	2665500	1	0.0340136054421	1887	0	0.02	0	11618	0
<input type="checkbox"/>	2815500	2817000	1	0.0340136054421	2459	0	0.0268425964592	0	-9280	0

Sine waves:

FILTER ROWS

<input type="checkbox"/>	start	end	n_periods	sine_freq	am_time_mean	am_amplitude_mean
<input type="checkbox"/>	6283500	6285000	8	238.21832941679128	0.01679131916420052	42337.375
<input type="checkbox"/>	6381000	6382500	8	248.9852595599231	0.016065208065208066	42190.625

СРАВНЕНИЕ ХАРАКТЕРИСТИК СИГНАЛОВ

	CS/AD7	CS/n	agn/AD7	agn/n
1 Количество импульсных песен	37	100	203	112
2 Среднее количество импульсов в песне	11.2	8.9	8.2	7
3 Количество синусных песен	2	17	19	20
4 Среднее время синусных песен (с)	0.5	0.5	0.7	0.4
5 Средняя энергия сигнала	13882.2	25381.1	36436	20323.4
6 Среднее количество периодов	78.5	80.8	102.9	53.5
7 Доля песен в 300 с записи (%)	2	6.5	11.1	5.6

Аннотирование
человеком

	CS/AD7	CS/n	agn/AD7	agn/n
1 Количество импульсных песен	25	132	128	90
2 Среднее количество импульсов в песне	2.8	2.7	2.7	2.6
5 Средняя энергия сигнала	14616.8	22579.7	29730.8	37425.8
3 Количество синусных песен	6	33	47	32
4 Среднее время синусных песен (с)	0.4	0.2	0.2	0.3
6 Среднее количество периодов	49.2	40.4	36.5	43.7
7 Доля песен в 300 с записи (%)	1.1	3.6	6.14	4.2

Аннотирование нашим
алгоритмом

БИОНACK 2018