

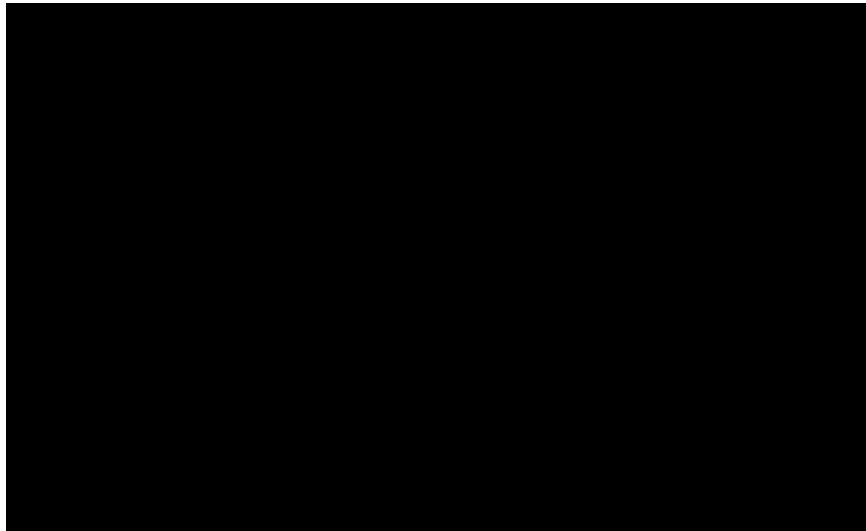
To: "Franck Chastagnol" <fchastagnol@youtube.com>
From: "Matthew Liu" <matthew@youtube.com>
Cc: "Frey Waid" <frey@youtube.com>, [REDACTED], "Steve Chen"
<steve@youtube.com>, "Chris Maxcy" <chris@youtube.com>
Bcc:
Sent Date: 2006-08-16 23:52:31 CST
Subject: Re: [Uncle] Result of fingerprinting on random videos

Guys,

I ran through all the videos one by one. Quite a pain but I classified the videos by type to the best of my knowledge (for a few of the videos I was unsure...there were also some videos that have been removed). If we try to separate between copyright (Music Video, Live Performance, TV shows/commercials, Karaoke, Mashups, Compilations) vs. non copyright (no melody or user-generated music) the breakdown is:

-Copyright - roughly 47% has some sort of copyright content in it (bits of music, anime, footage from tv commercials, etc...not all of it is necessarily infringing)

-Non copyright - 50%



Thanks,
Matt

Attachments:

ATT00601.htm
result_random_video Matt.xls

DATE: 8.22.08 EXHIBIT# 8
DEPONENT: MAXCY, CHRIS

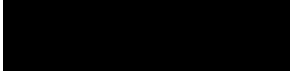
CASE: Viacom, et al., v. YouTube, et al., The Football
Association Premier League, et al., v. YouTube, et al.,
Case Nos. 07-CV-2203 and 07-CV-3582
A. Ignacio Howard, CLR, RPR, CSR No. 9830

Highly Confidential

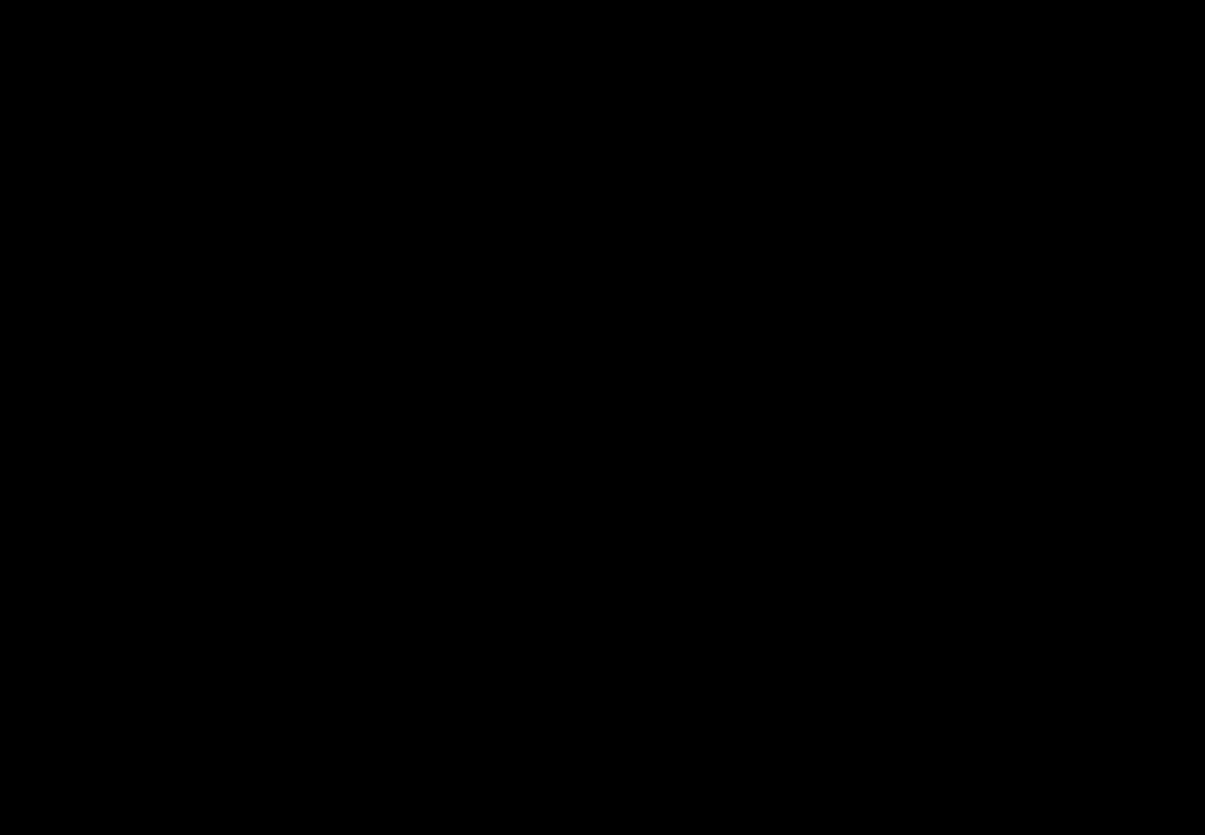
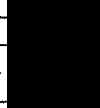
000001-00730943

GraceNote (GN) vs AudibleMagic (AM) accuracy tests

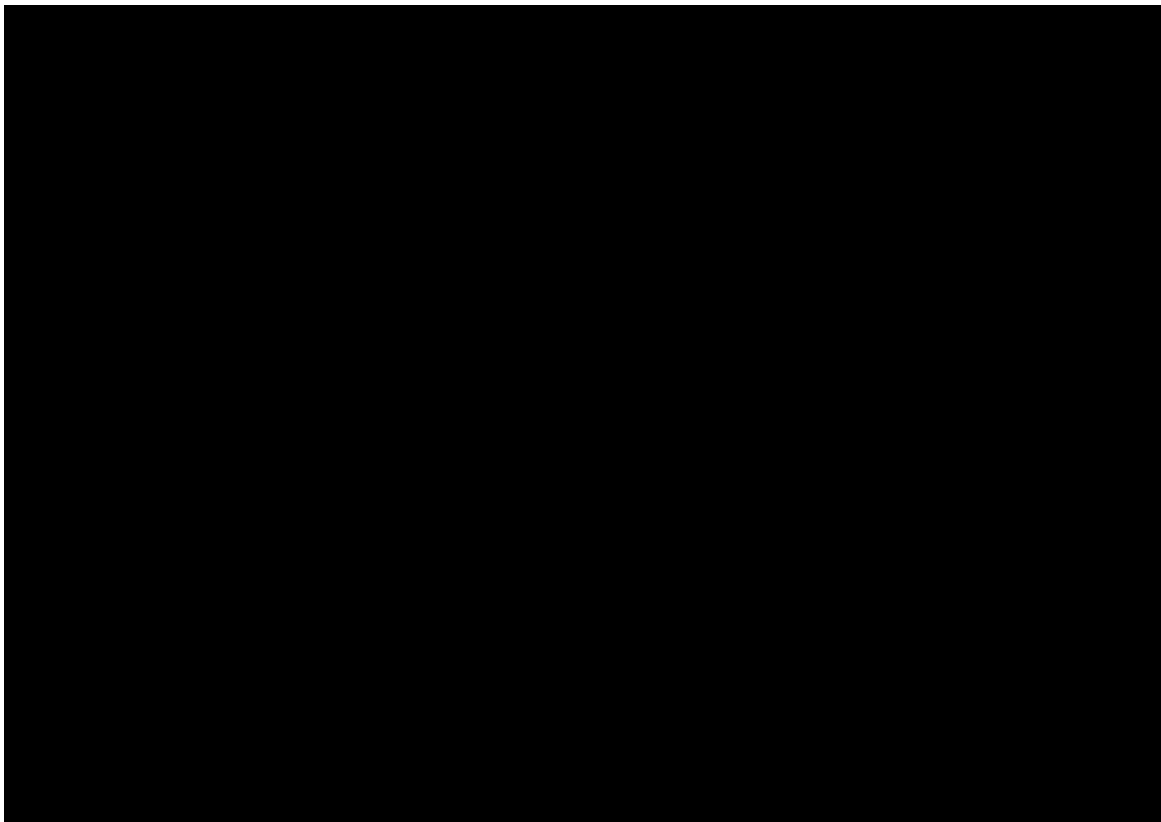
Date: Sun, 13 Aug 2006 23:27:29



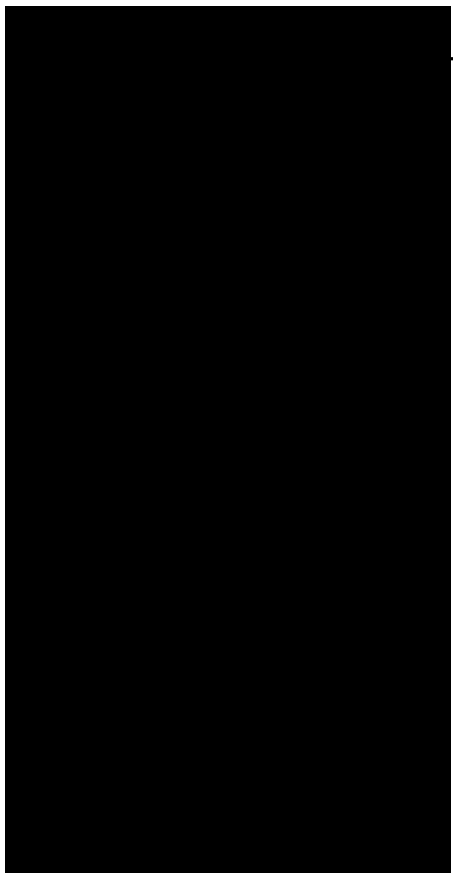
Nb videos	
GN matches	
AM matches (*)	

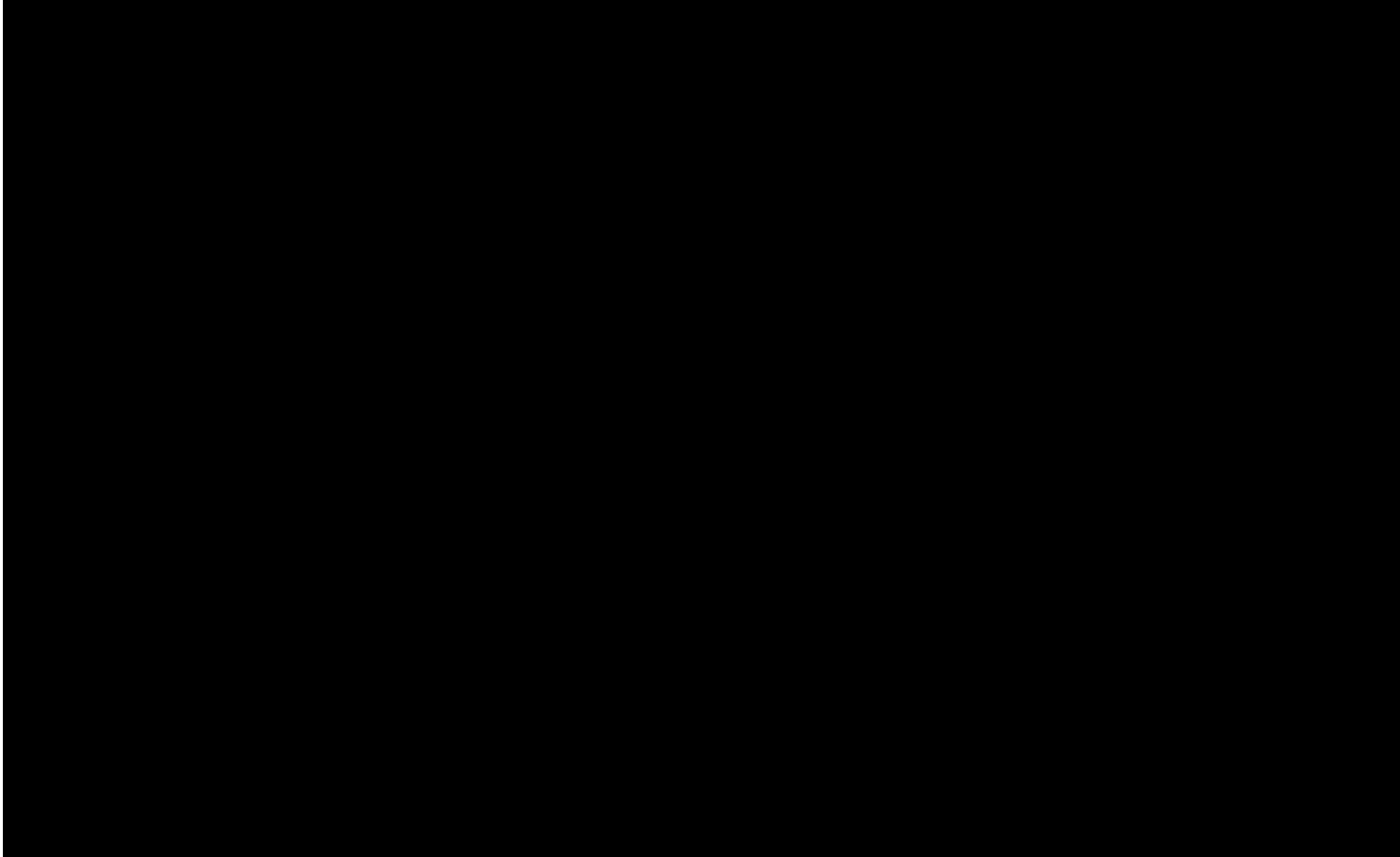


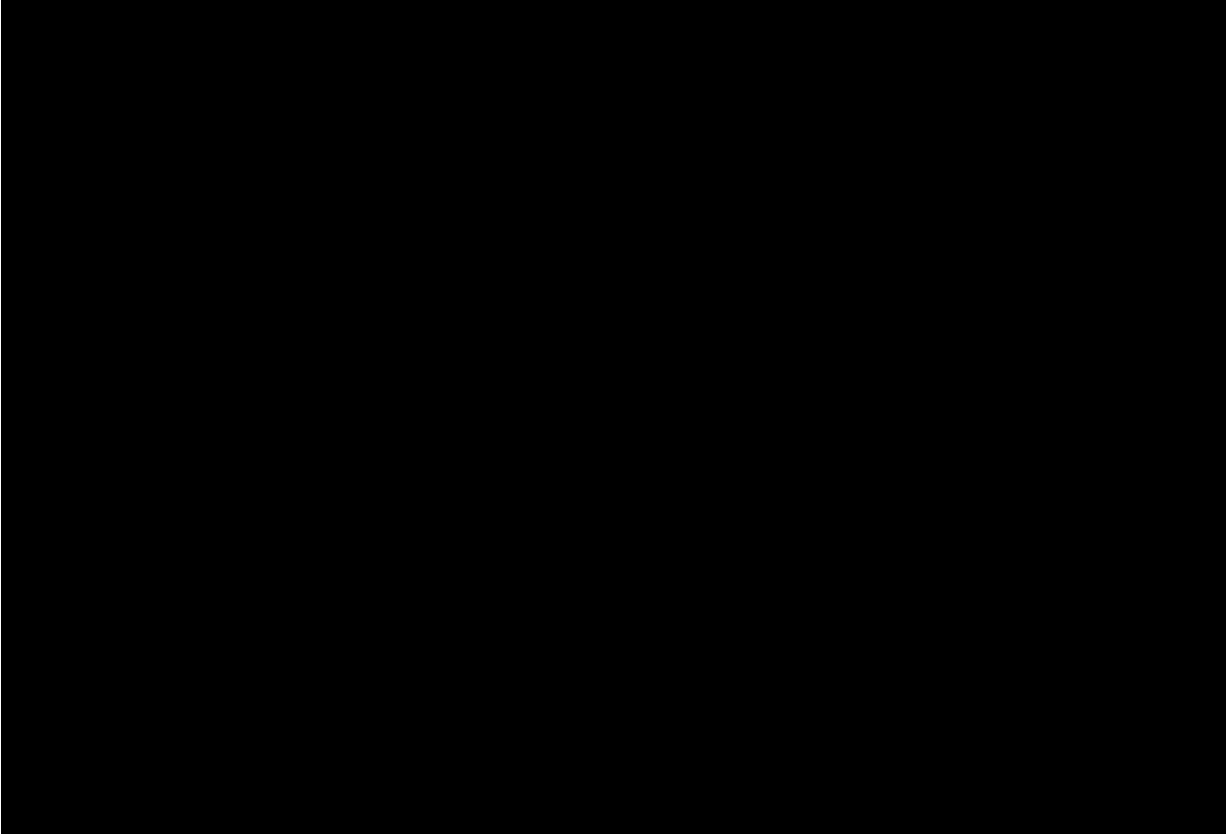
	Ratio



(*) 3 audio matched by GN could not be run through AM due to some issue with snippets







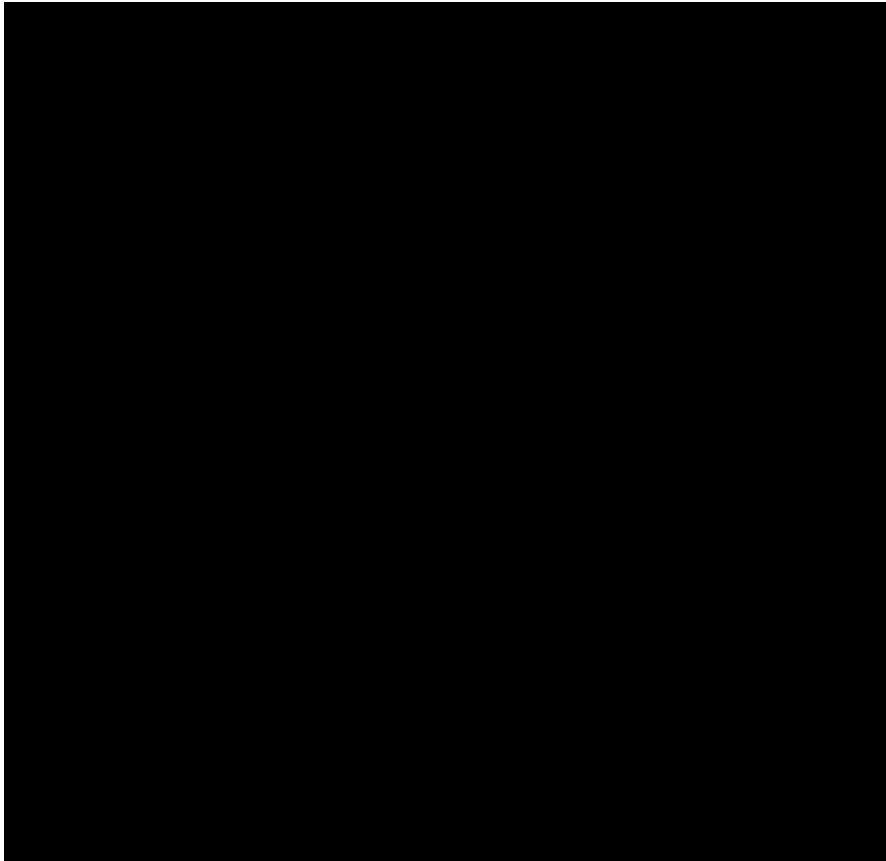
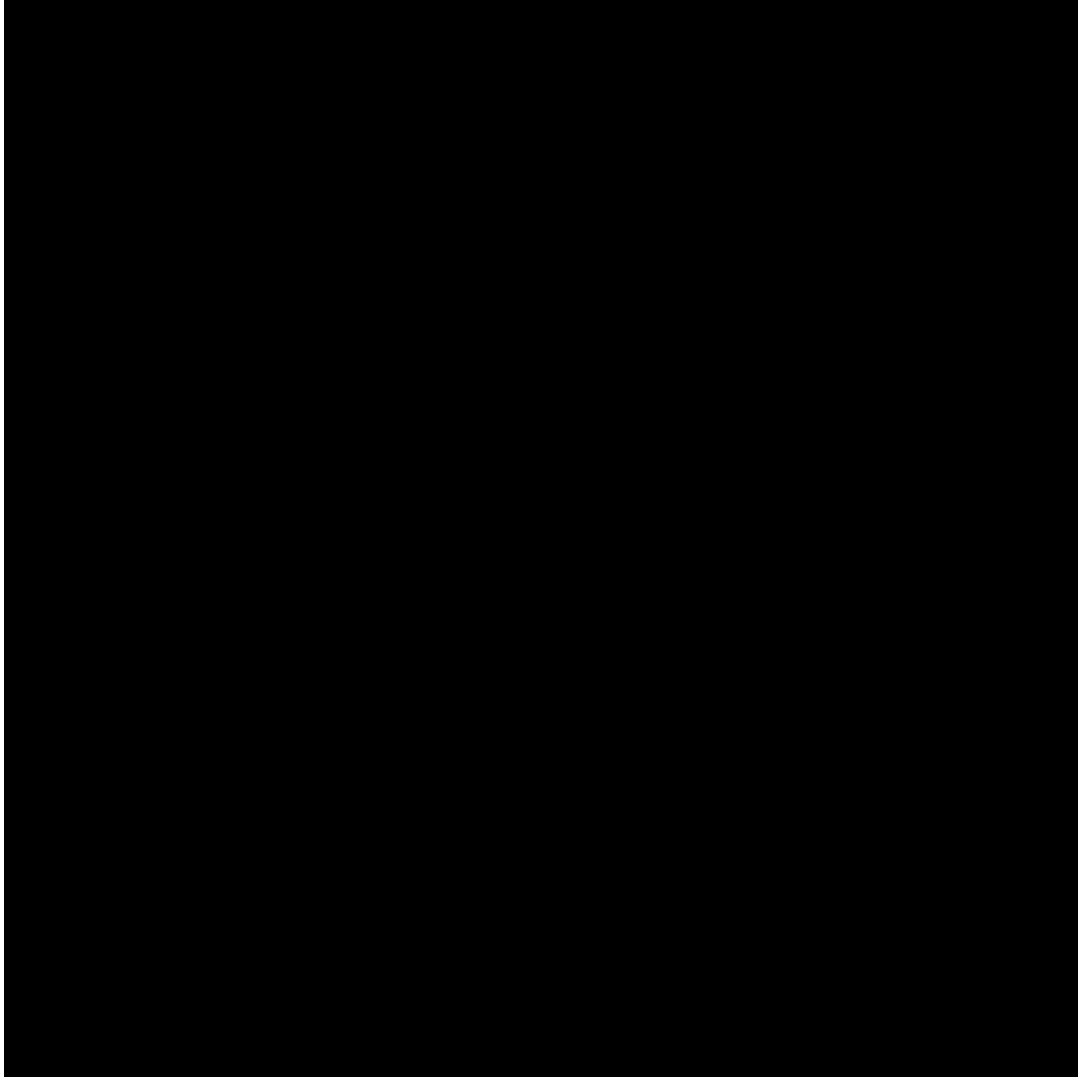
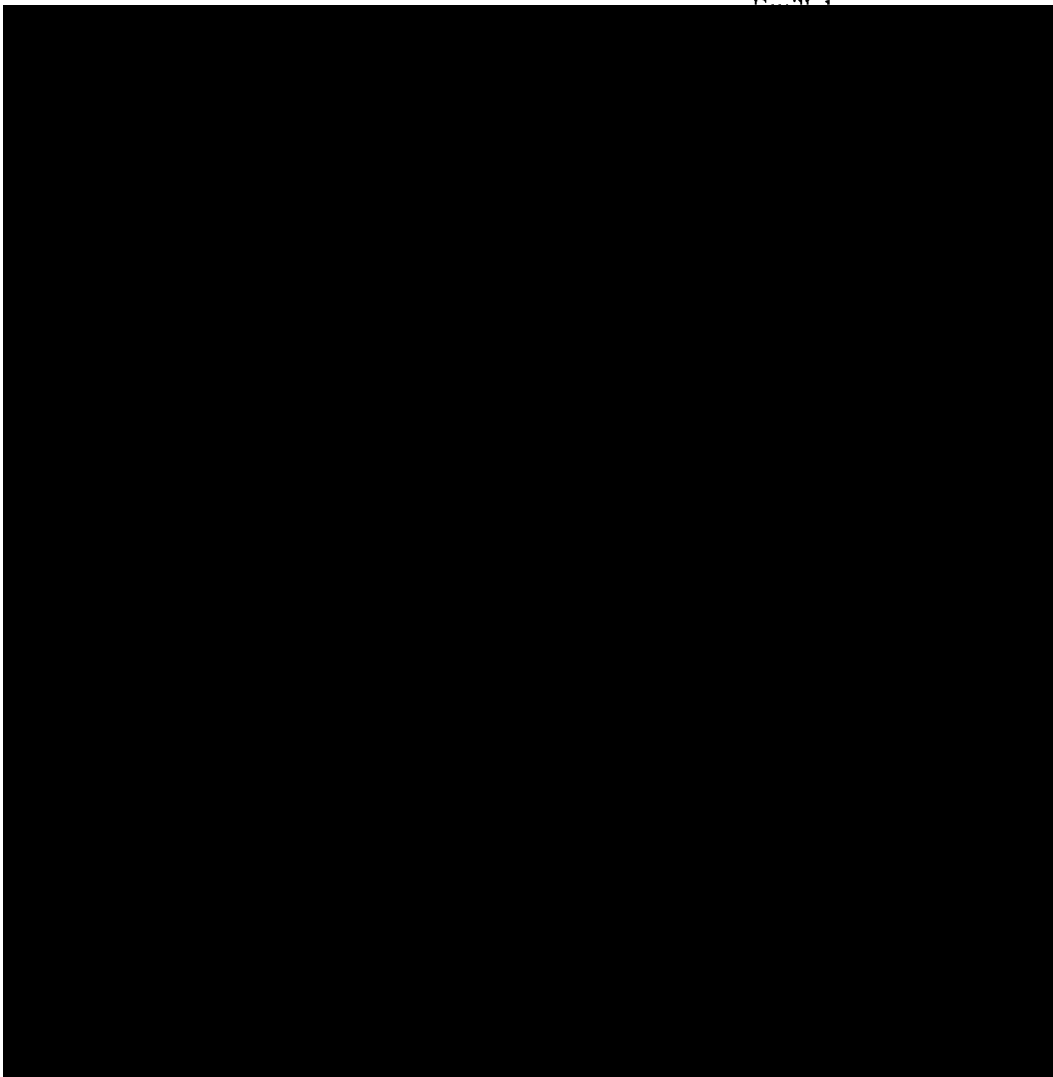


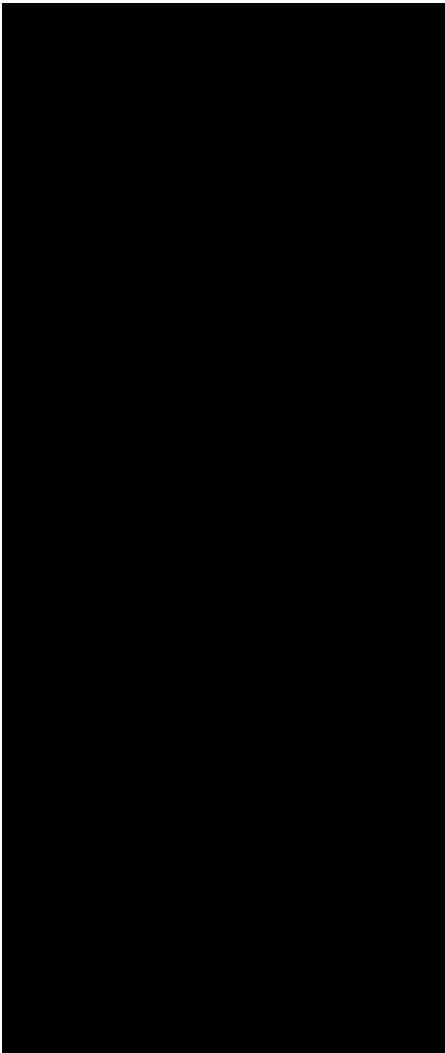
Exhibit 1

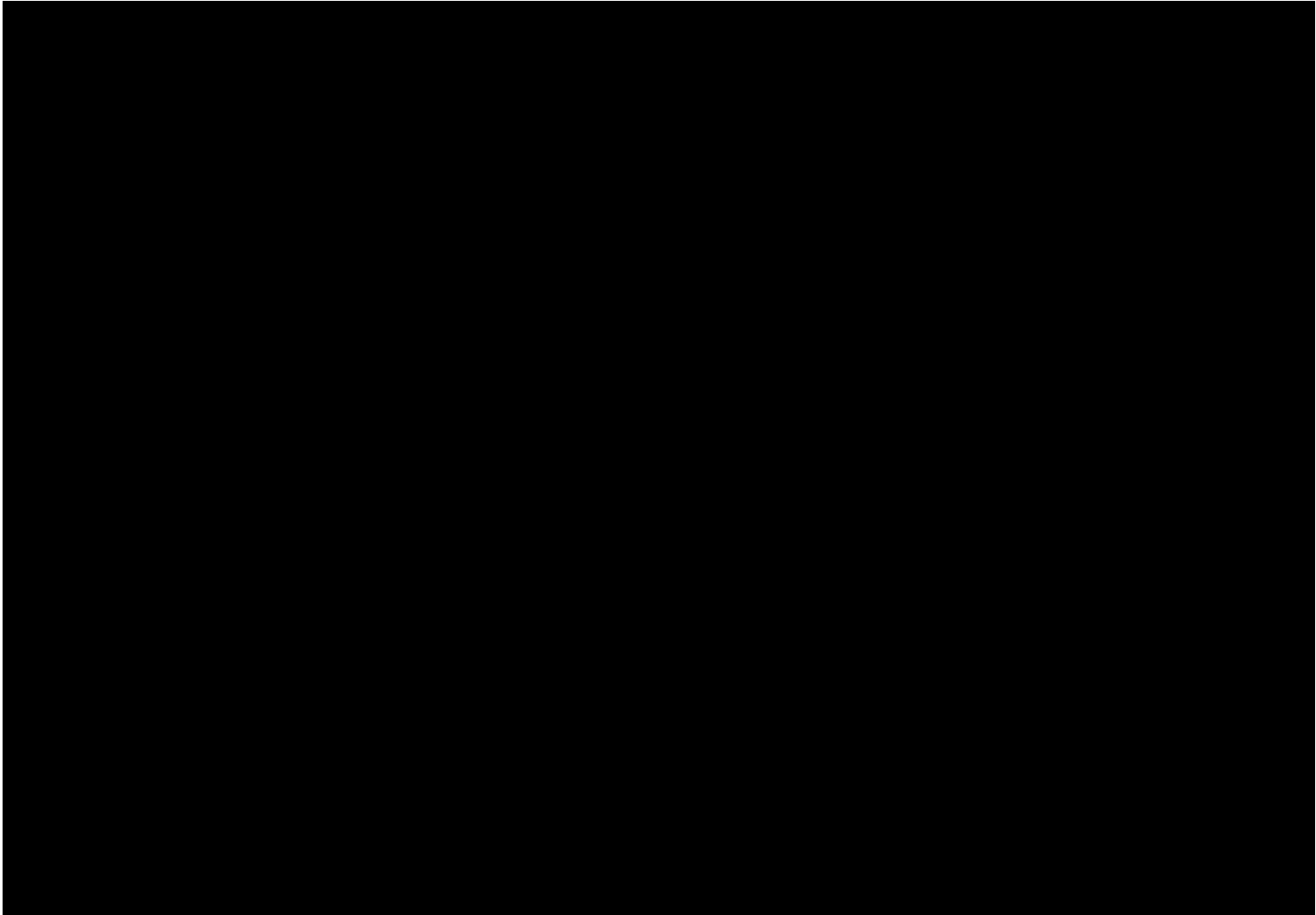


Highly Confidential

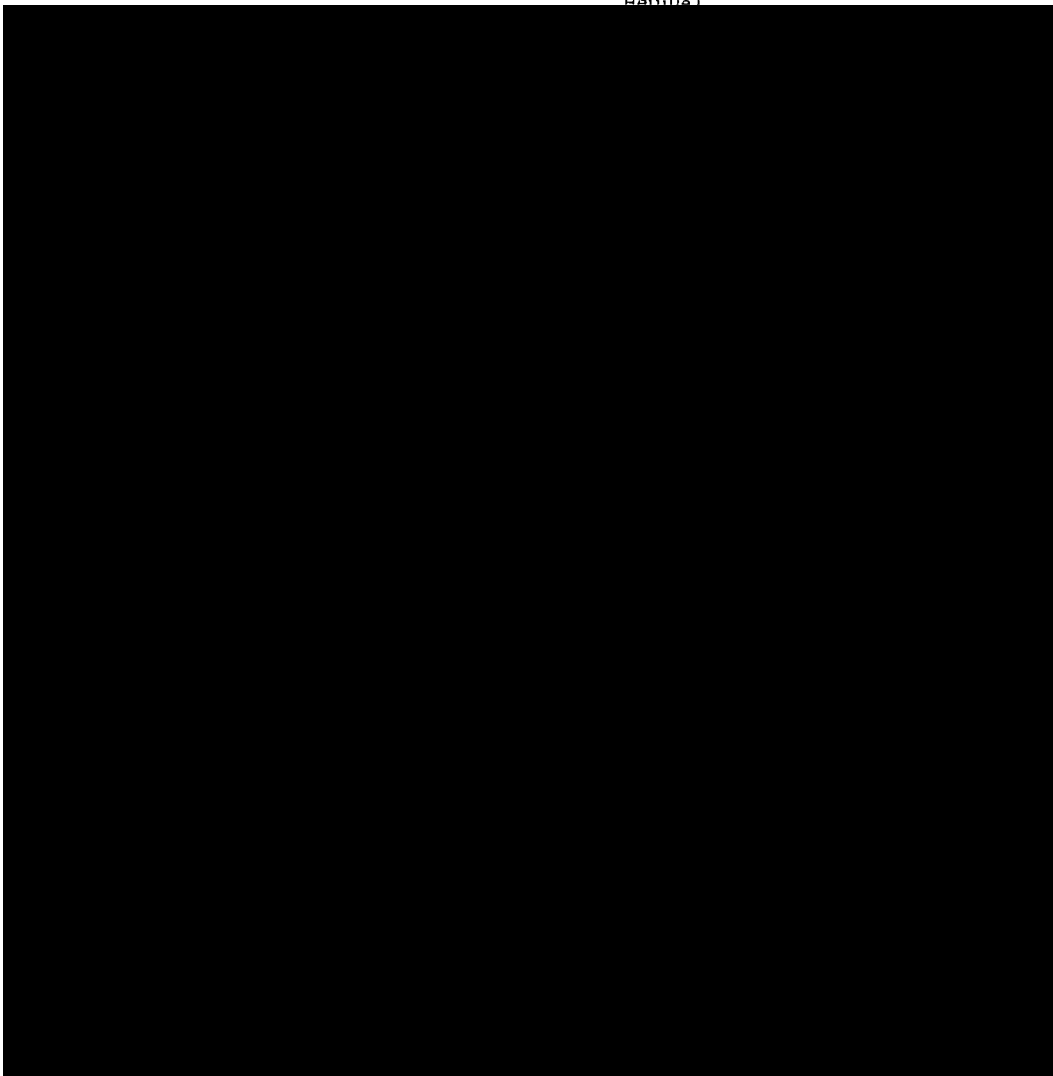
EO 13526-00730951

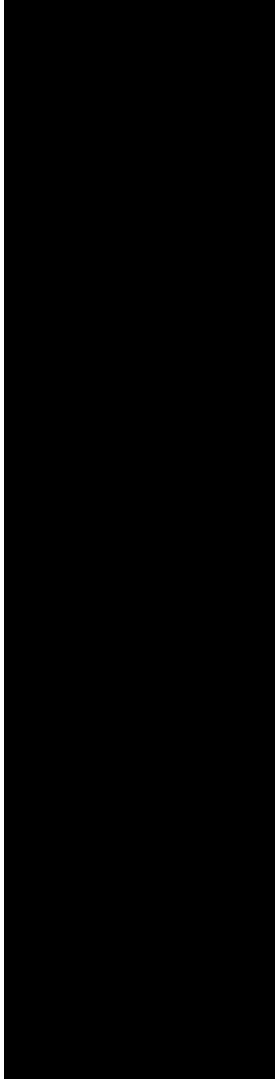




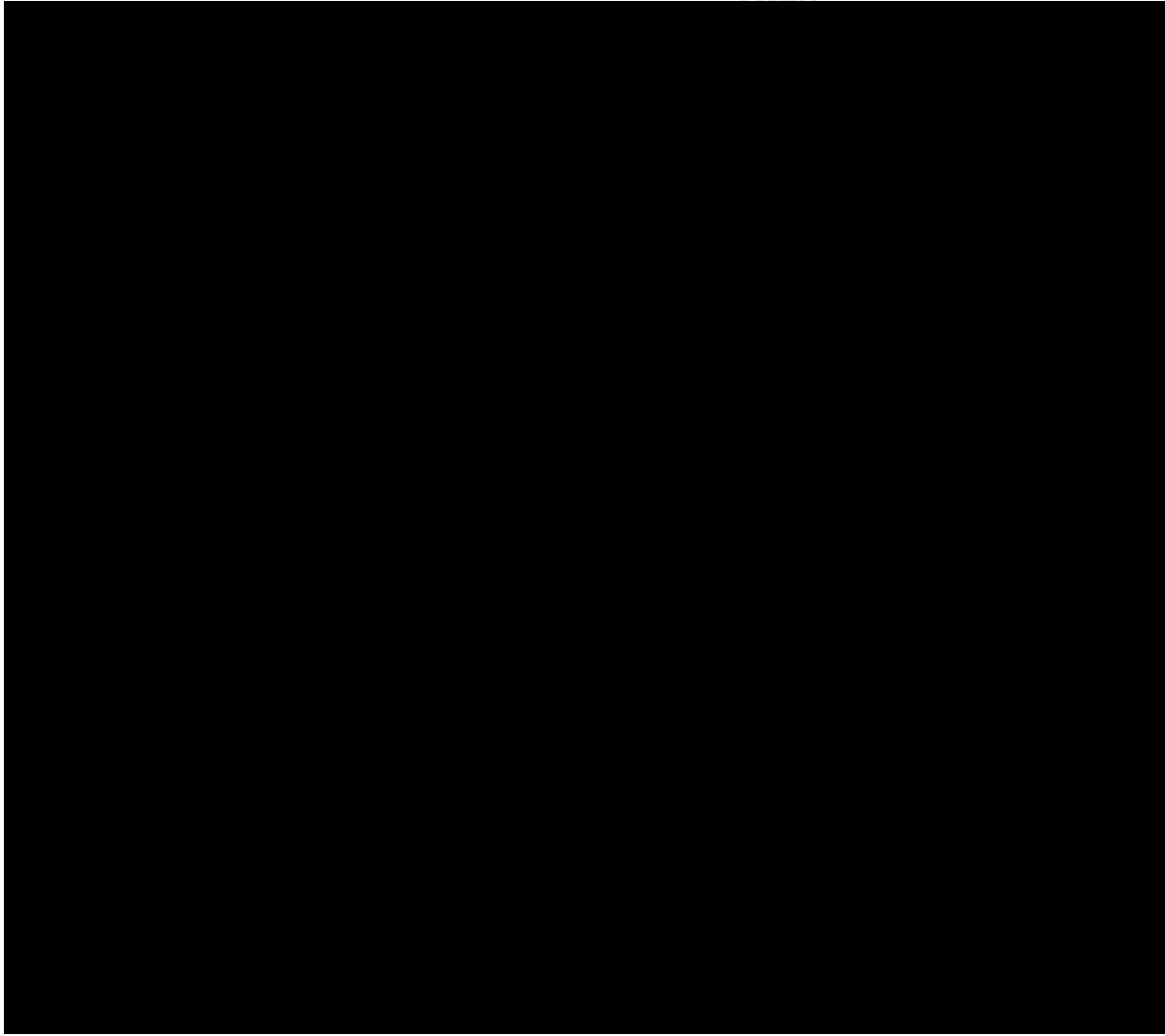


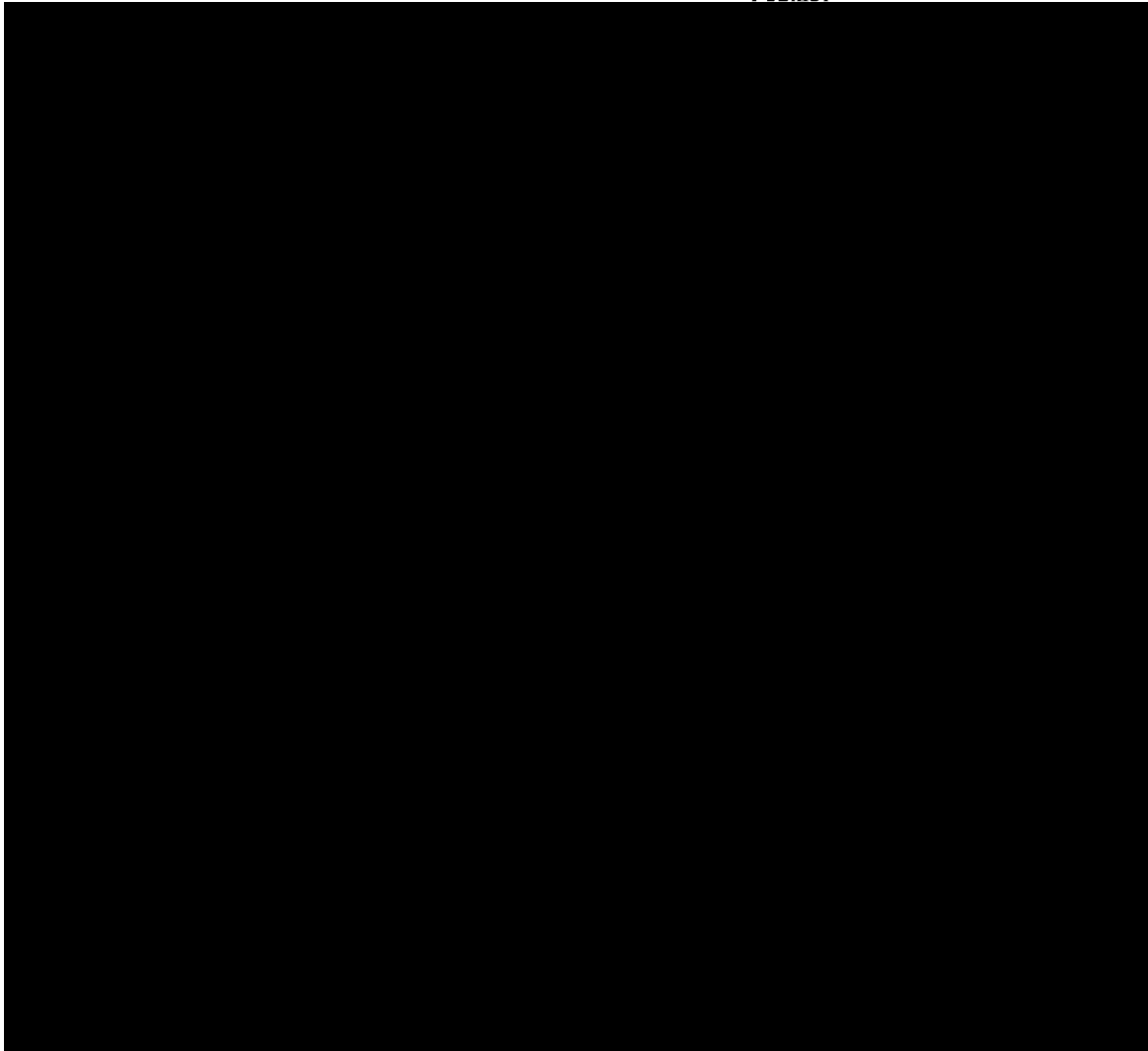
Enville1

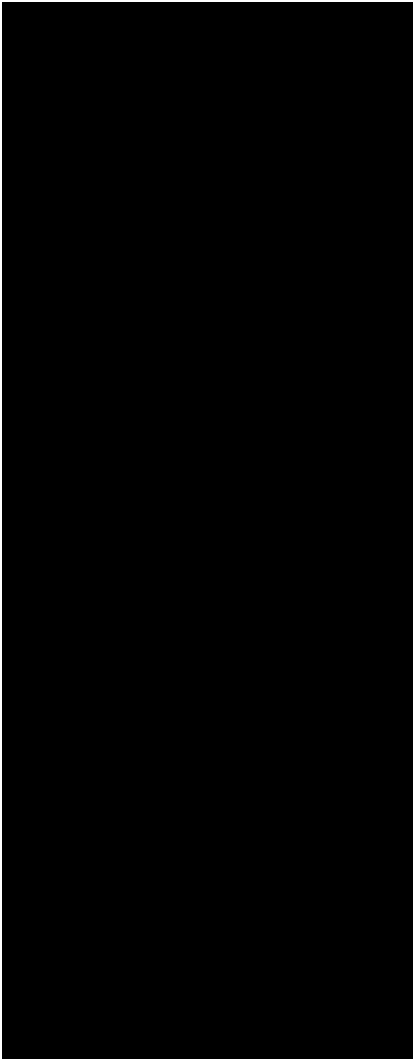


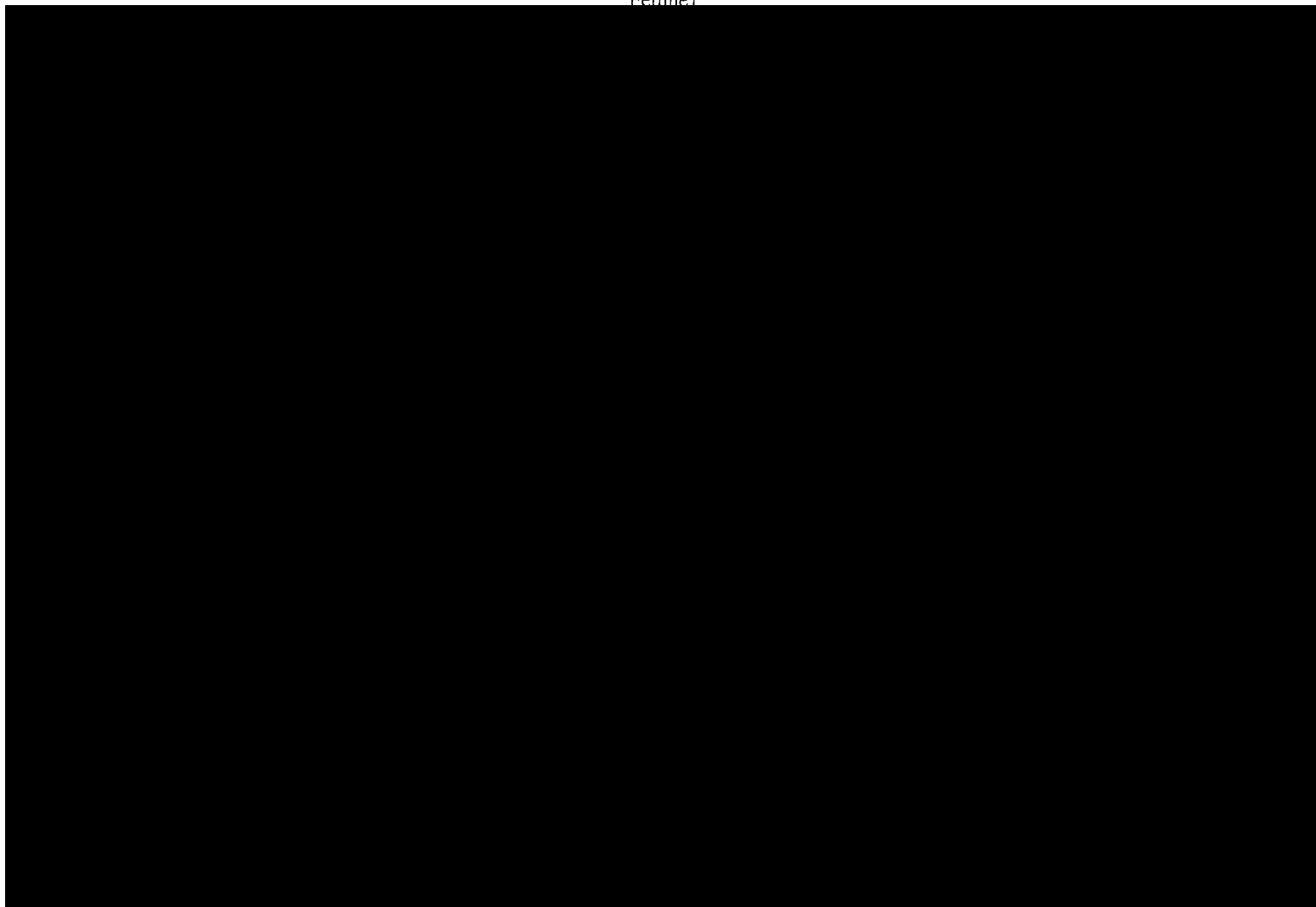


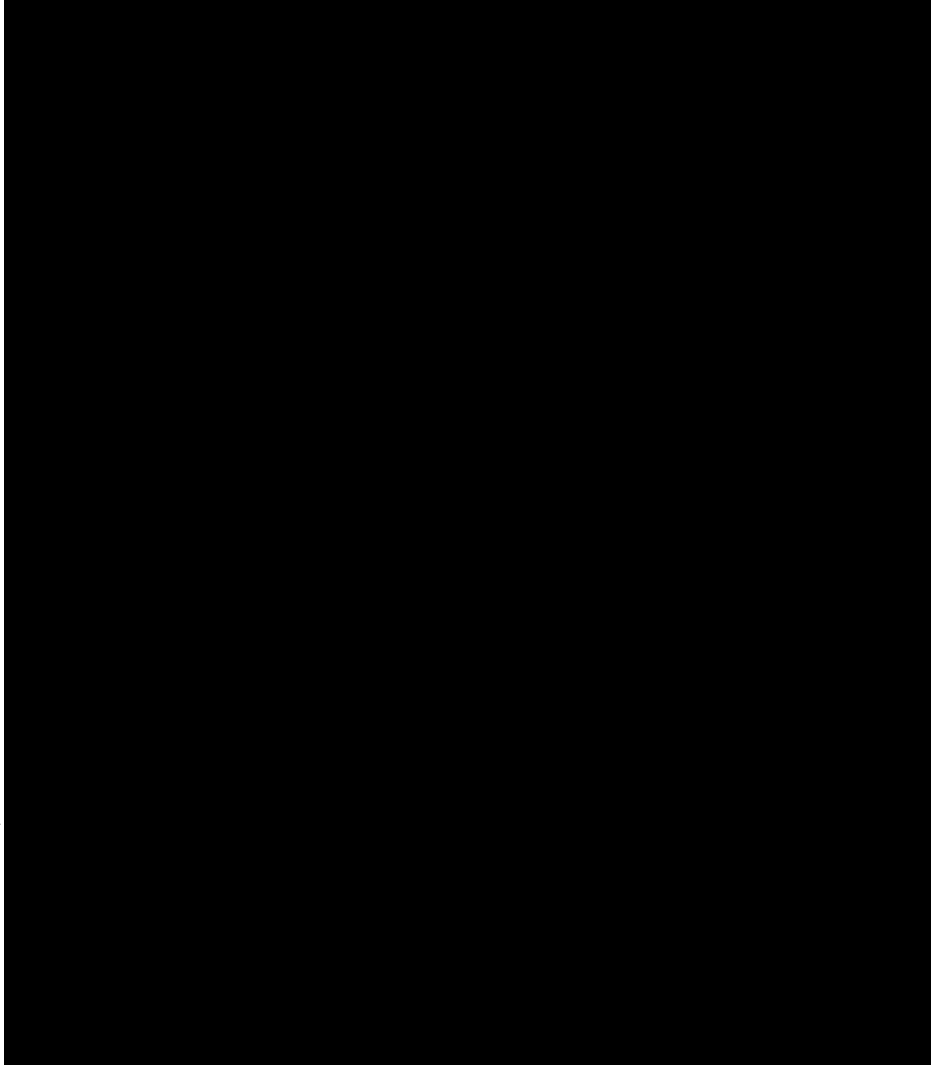
Feuille1

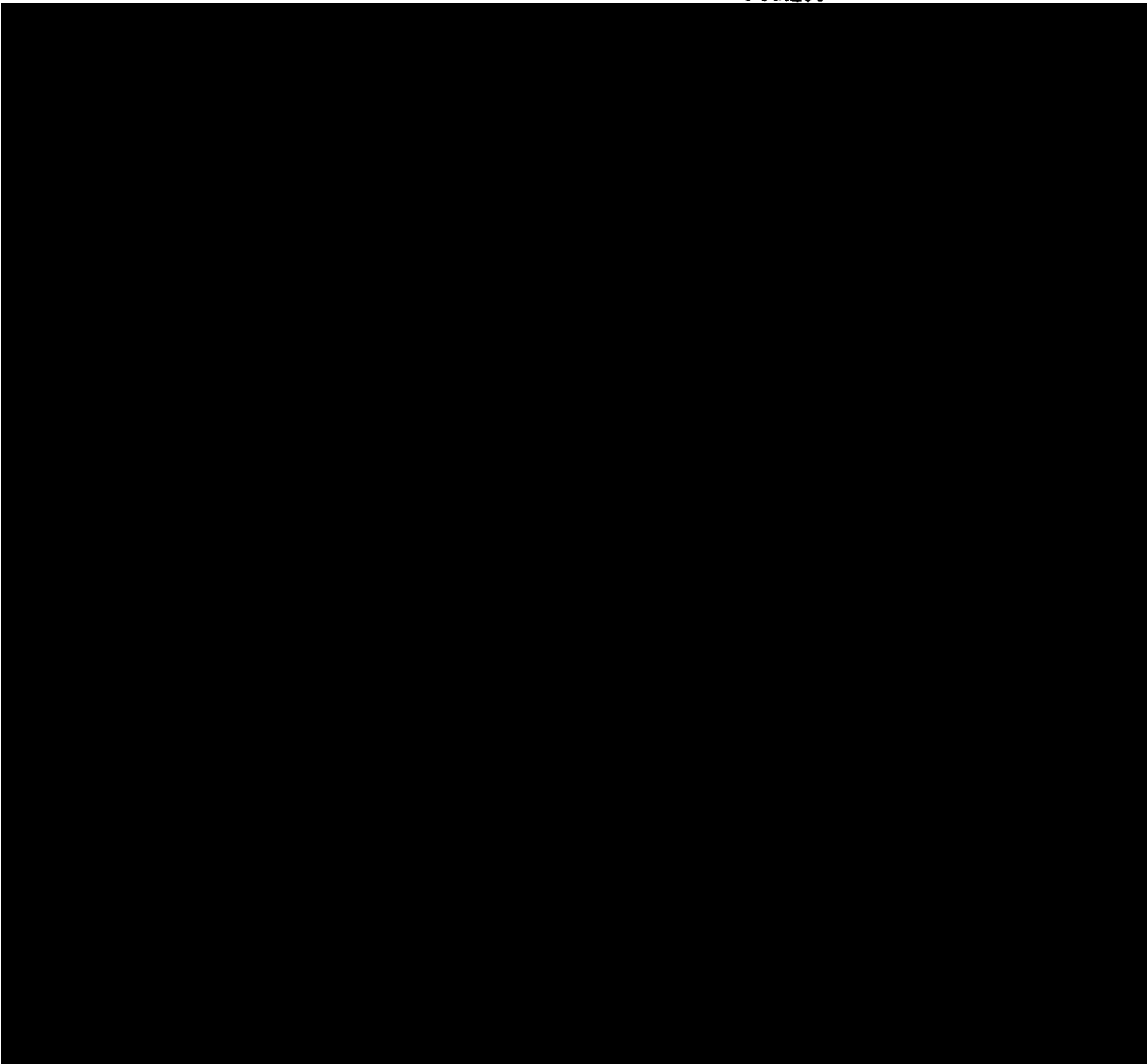


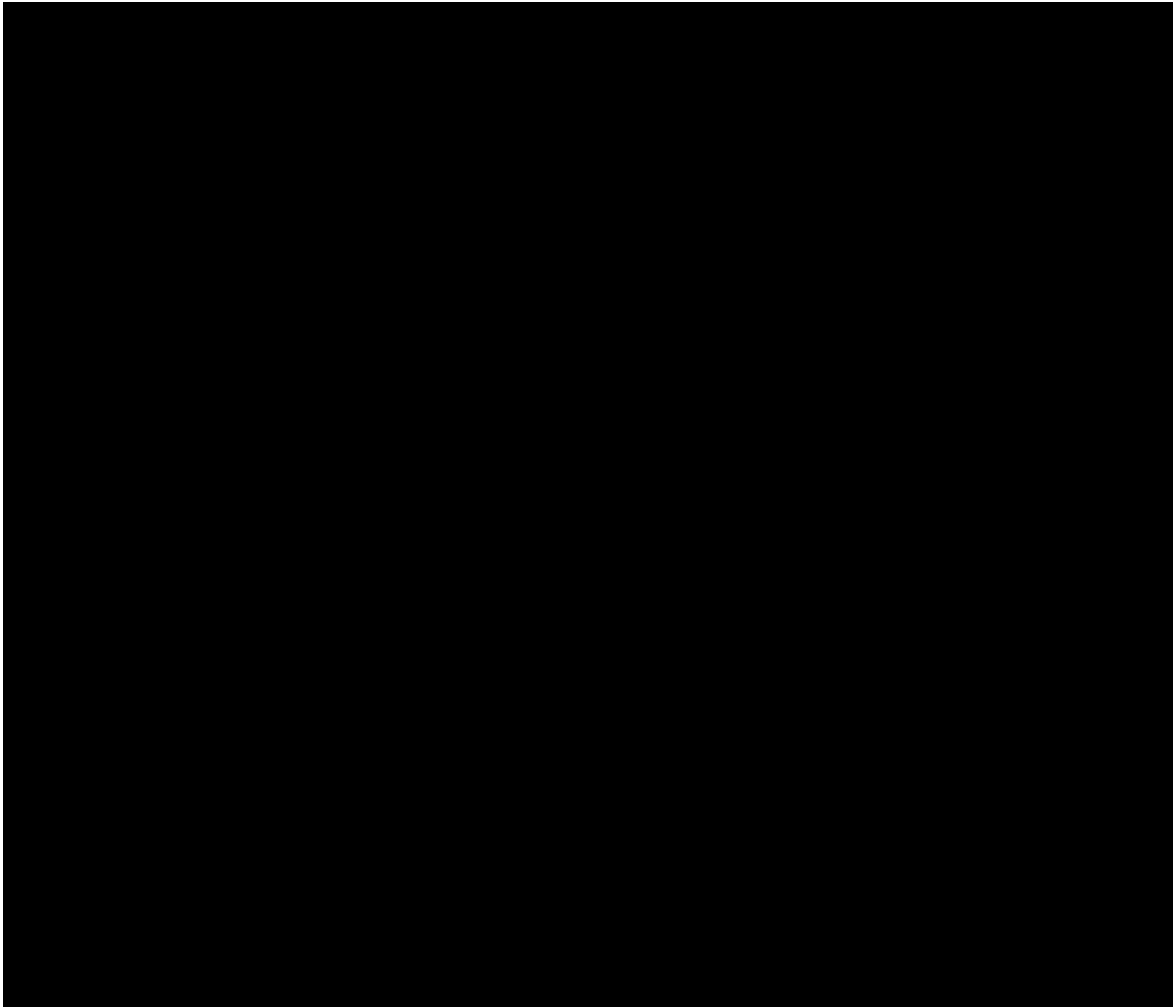


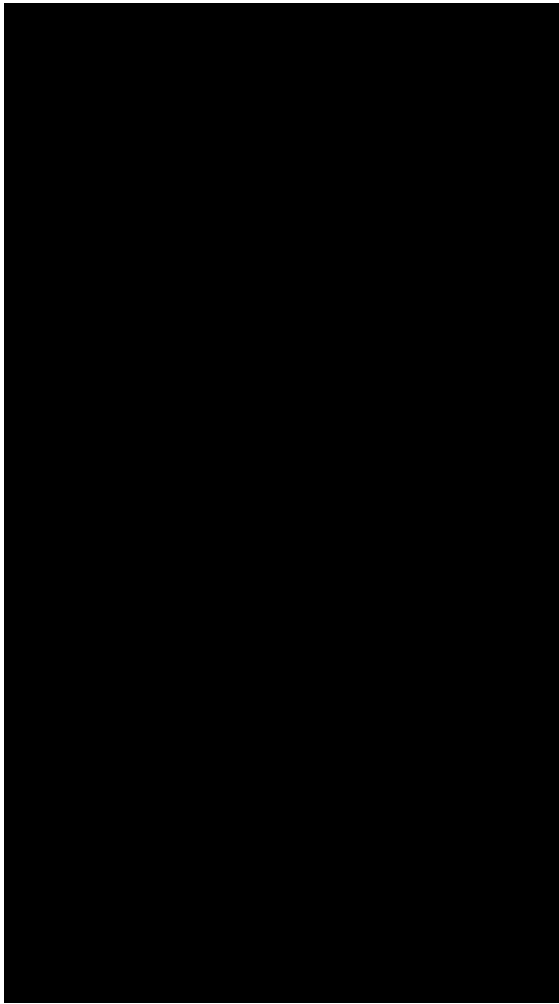


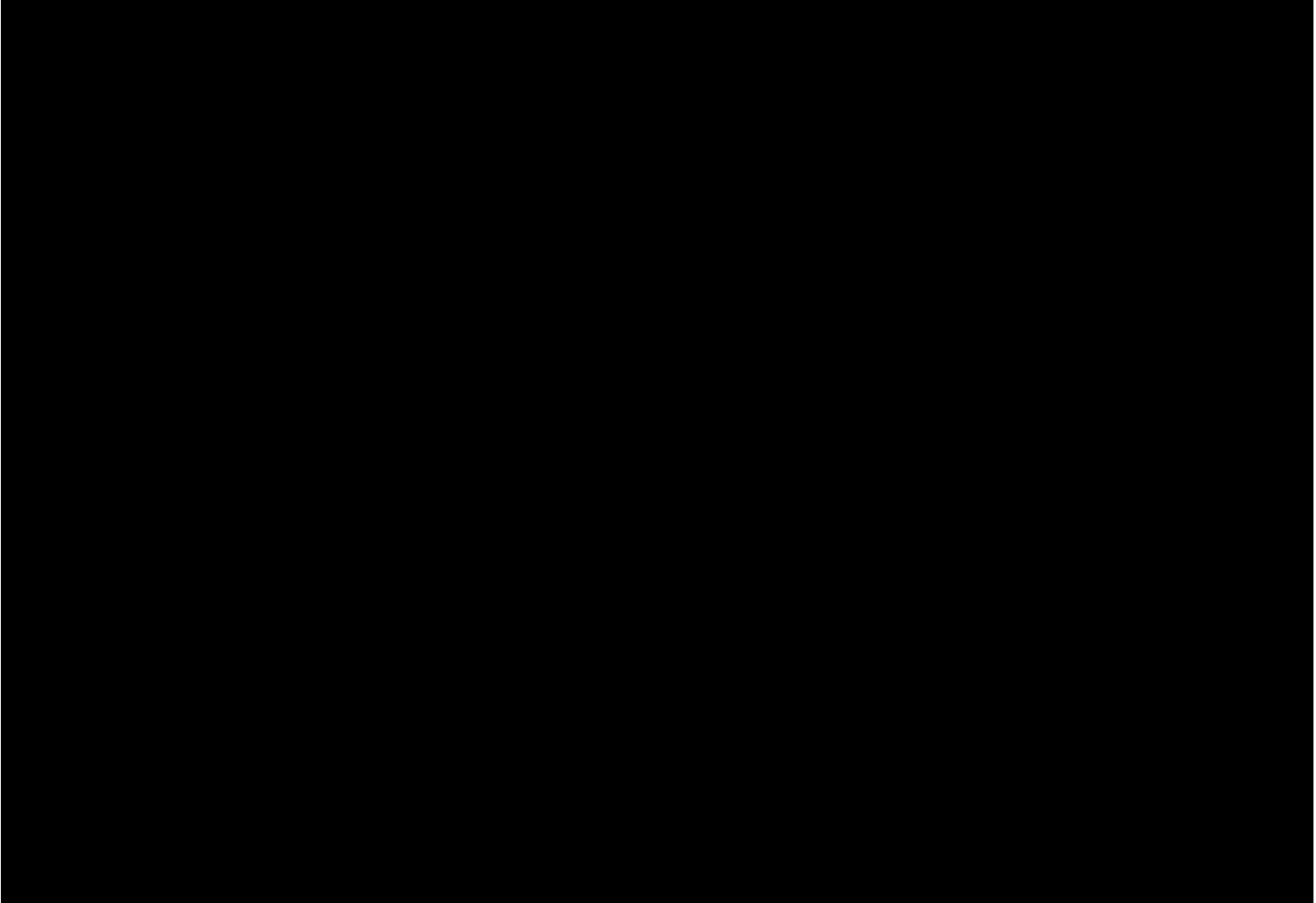


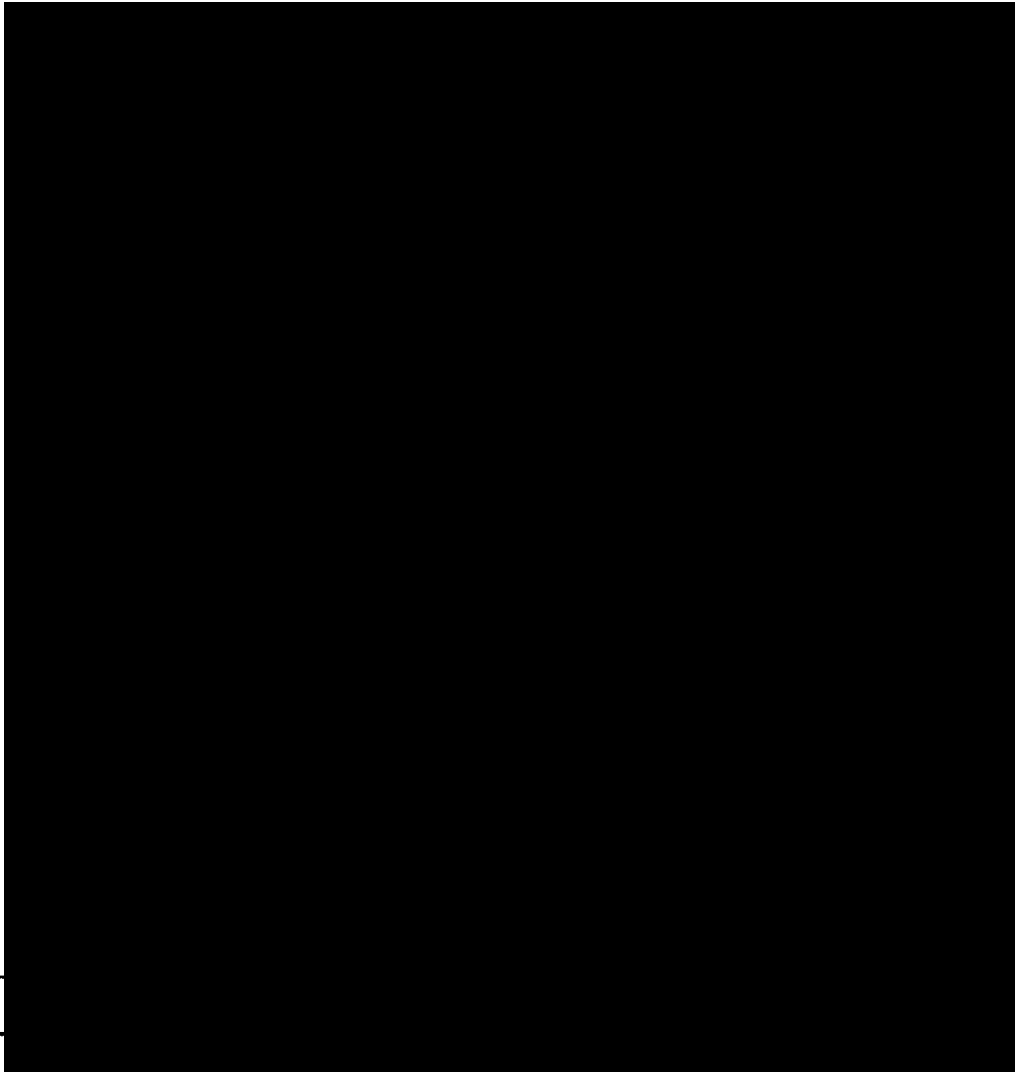




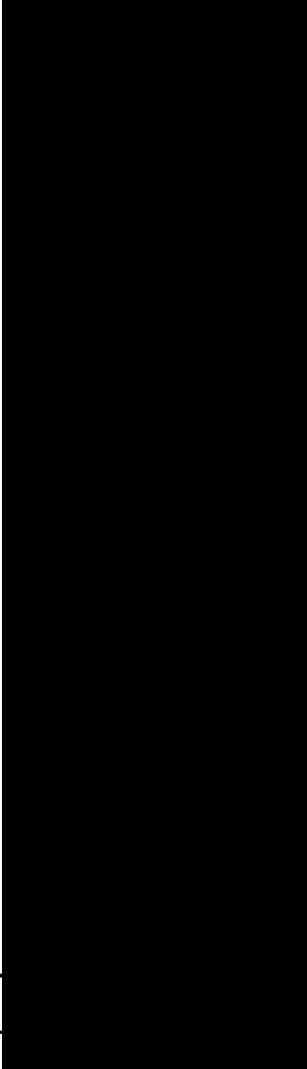








Feuille1



Copyrighted
Non-copyrighted

Feuillel
46.97%
50.00%

Highly Confidential

Feuille1

600001-00730970

Page 27

Feuille1

Page 29

Highly Confidential

600001-00730972

Highly Confidential

Feuille1

600001-00730973

Page 30

On Aug 14, 2006, at 10:02 AM, Frey Waid wrote:

=BLOCKQUOTE type="cite">Hi Franck,

Out of curiosity, did any of the XML responses from either GN or AM contain distributor info -- Warner or otherwise? Maybe, we could submit something we know is Warner content (perhaps private music file) to see if we get the distributor name in the feed. No, neither AM nor GN seem to send us back this info, even though the XML response has provision for returning it. I need to follow up with them to understand why they do not populate these fields in the response.

It does seem clear to me that we will want to investigate video fingerprinting technology so that we can differentiate between copyrighted audio tracks and copyrighted music videos. agreed, video fingerprinting would bring us to the next level.

Also, I'm not sure we have a good solution yet for identifying live concerts: descriptive text filtering will almost certainly be insufficient. AM does not do a good job at identifying music on live performance. GN is better. Now are you saying we need to differentiate copyrighted song from a live performance vs copyrighted song from a music video or a TV show? Let's talk about that - I may be missing something.

Thanks, Franck

Franck Chastagnol wrote: Hi team,

I ran both GraceNote and AudibleMagic against a random set of 133 videos from our site: - GN identified copyrighted music in [REDACTED] of these videos. - AM identified copyrighted music in [REDACTED] of these videos.

This seems to confirm the results of the previous test in terms of matching technology: GN is superior.

I did not have time to look at all the videos with a match but it seems that in most cases, these are personal videos that users edited to add music to. I found only a couple of music video clips.

Now, one of the conclusion that I think we should also draw from these tests is that it seems we have a pretty high percentage of our content that will be flagged as copyrighted as soon as we start using fingerprinting technology. Note that initially we will fingerprint only against Warner catalog so the percentage of match will certainly be lower. But as we start signing up new content partners, it will increase.

All the data is available

at: [https://trac.sjl.youtube.com/trac/attachment/wiki/FingerprintEval/SPAN class="Apple-converted-space"> result_random_video.xls](https://trac.sjl.youtube.com/trac/attachment/wiki/FingerprintEval/SPAN%20class%3D%22Apple-converted-space%22%3Eresult_random_video.xls)

Let me know any question, Franck ---

Uncle mailing list Uncle@youtube.com <https://dev.youtube.com/mailman/listinfo/uncle>

Uncle mailing list Uncle@youtube.com <https://dev.youtube.com/mailman/listinfo/uncle>
=SPAN class="Apple-style-span" style="border-collapse: separate; border-spacing: 0px 0px; color: rgb(0, 0, 0); font-family: Helvetica; font-size: 12px; font-style: normal; font-variant:

normal; font-weight: normal; letter-spacing: normal; line-height: normal; text-align: auto; =khtml-
text-decorations-in-effect: none; text-indent: 0px; =apple-text-size-adjust: auto; text-transform:
none; orphans: 2; =hite-space: normal; widows: 2; word-spacing: 0px; ">Matthew =iuProduct
Manager510.921.6684 | matthew@youtube.com

=/DIV>