



UNIVERSITY of MARYLAND
MEDICAL CENTER



ChimerAnalyzer

An Automated Approach to Streamline
Bone Marrow Engraftment Analysis

Jon Young, MD
PGY3 – Pathology

Bone Marrow Engraftment (BME) Analysis

- ▶ Hematopoietic stem cell transplantation is a common therapy for hematologic malignancies
- ▶ BME analysis used to evaluate the engraftment of donor bone marrow within transplant recipients
- ▶ BME analysis performed to monitor transplant outcomes by confirming successful engraftment and monitoring for relapse

Bone Marrow Engraftment Analysis

- ▶ Analyzing short tandem repeat (STR) polymorphisms in pre-transplant and post-transplant specimens is the current clinical standard of care
- ▶ STR polymorphisms can be used to differentiate between patient and donor DNA as the number of repeat units at a STR locus varies between individuals
- ▶ Compare the relative amounts of donor alleles to patient alleles, as determined by multiplex PCR and capillary electrophoresis
- ▶ Calculate % chimerism

Pre-Transplant

- ▶ Determine patient and donor STR pattern/genotype/fingerprint
- ▶ Patient and donor STR patterns are analyzed to identify unshared alleles, which can be used to differentiate between the two in follow-up analysis
- ▶ ‘Informative’ loci

16 different STR markers (tetranucleotide repeats)

Blue	D8S1179	(128-168 bp)		
	D21S11	(189-243 bp)		
	D7S820	(258-294 bp)		
	CSF1PO (5)	(306-342 bp)		
Green	D3S1358	(114-142 bp)		
	THO1 (11)	(169-189 bp)		
	D13S317	(217-245 bp)		
	D16S539	264-284		
	D2S1338	311-355		
Yellow	D19S433	113-131		
	VWA (12)	(157-197 bp)		
	TPOX (2)	(225-249 bp)		
	D18S51	(273-341 bp)		
Red	Amelogenin	106,112		
	D5S818	(135-171 bp)		
	FGA (4)	(219-267 bp)		

Sample Electrophoresis Data

		Dye/Sample Peak	Size	Height	Area
26		G,1	59.38	102	1197
27		G,2	119.39	435	5860
28		G,3	123.35	7479	96912
29		G,4	132.18	83	931
30		G,5	134.86	166	2098
31		G,6	136.36	95	1104
		Dye/Sample Peak	Size	Height	Area
32		G,7	165.17	200	2543
33		G,8	166.89	4001	39129
34		G,9	169.82	88	1097
35		G,10	170.96	4082	39854
36		G,11	211.81	89	437
37		G,12	217.66	90	580
38		G,13	228.22	451	4970
39		G,14	232.18	8137	95763
40		G,15	259.97	85	894
41		G,16	263.9	4040	45791
42		G,17	270.98	84	418
43		G,18	271.97	198	2235
44		G,19	276.03	3755	43782
45		G,20	306.98	149	1644
46		G,21	311.3	2894	34905
47		G,22	315.57	148	1801
48		G,23	319.81	2528	30151
49		G,24	321.51	118	741
50		G,25	325.61	114	631
51	Yellow	Y,1	109.4	209	2129
52	Yellow	Y,2	113.33	3574	32904
53	Yellow	Y,3	119.25	512	4976
54	Yellow	Y,4	123.21	3523	34724
55	Yellow	Y,5	134.79	123	1396

Pre-Transplant Genotype

Patient

Blue	D8S1179	(128-168 bp)		
	D21S11	(189-243 bp)		
	D7S820	(258-294 bp)		
	CSF1PO (5)	(306-342 bp)		
Green	D3S1358	(114-142 bp)	123	123
	TH01 (11)	(169-189 bp)	167	171
	D13S317	(217-245 bp)	232	232
	D16S539	264-284	264	276
	D2S1338	311-355	311	320
Yellow	D19S433	113-131	113	123
	VWA (12)	(157-197 bp)		
	TPOX (2)	(225-249 bp)		
	D18S51	(273-341 bp)		
Red	Amelogenin	106,112		
	D5S818	(135-171 bp)		
	FGA (4)	(219-267 bp)		

Peak heights > 500 represent an allele...

Not all peaks >500 represent alleles...

63	Y,13	221.73	3248	37834
64	Y,14	232.18	1684	12398
65	Y,15	233.58	3144	37753
66	Y,16	263.9	411	4699
67	Y,17	276.03	343	3900
68	Y,18	278.02	162	1744

Blue	D8S1179	(128-168 bp)
	D21S11	(189-243 bp)
	D7S820	(258-294 bp)
	CSF1PO (5)	(306-342 bp)
Green	D3S1358	(114-142 bp)
	TH01(11)	(169-189 bp)
	D13S317	(217-245 bp)
	D16S539	264-284
	D2S1338	311-355
Yellow	D19S433	113-131
	VWA (12)	(157-197 bp)
	TPOX (2)	(225-249 bp)
	D18S51	(273-341 bp)
Red	Amelogenin	106,112
	D5S818	(135-171 bp)
	FGA (4)	(219-267 bp)

Not all peaks >500 represent alleles...

	Dye/Sample Peak	Size	Height	Area
32	G,7	165.17	200	2543
33	G,8	166.89	4001	39129
34	G,9	169.82	88	1097
35	G,10	170.96	4082	39854
36	G,11	211.81	89	437
37	G,12	217.66	90	580
38	G,13	228.22	451	4970
39	G,14	232.18	8137	95763
40	G,15	259.97	85	894
41	G,16	263.9	4040	45791
42	G,17	270.98	84	418
43	G,18	271.97	198	2235
44	G,19	276.03	3755	43782
45	G,20	306.98	149	1644
46	G,21	311.3	2894	34905
47	G,22	315.57	148	1801
48	G,23	319.81	2528	30151
49	G,24	321.51	118	741
50	G,25	325.61	114	631
51	Y,1	109.4	209	2129
52	Y,2	113.33	3574	32904
53	Y,3	119.25	512	4976
54	Y,4	123.21	3523	34724
55	Y,5	134.79	123	1396
63	Y,13	221.73	3248	37834
64	Y,14	232.18	1684	12398
65	Y,15	233.58	3144	37753
66	Y,16	263.9	411	4699
67	Y,17	276.03	343	3900
68	Y,18	278.02	162	1744

Blue	D8S1179	(128-168 bp)
	D21S11	(189-243 bp)
	D7S820	(258-294 bp)
	CSF1PO (5)	(306-342 bp)
Green	D3S1358	(114-142 bp)
	THO1 (11)	(169-189 bp)
	D13S317	(217-245 bp)
	D16S539	264-284
	D2S1338	311-355
Yellow	D19S433	113-131
	VWA (12)	(157-197 bp)
	TPOX (2)	(225-249 bp)
	D18S51	(273-341 bp)
Red	Amelogenin	106,112
	D5S818	(135-171 bp)
	FGA (4)	(219-267 bp)

'Informative'

- ▶ Patient relative to donor
- ▶ Donor relative to patient
- ▶ Criteria:
 - Unique alleles not shared by donor
 - Numerator not in a position of stutter down relative to donor
 - No overlap
 - No offscale
- ▶ Stop when 2 or 3 informative loci found

Equations

$$\frac{\text{Pt allele (peak height)}}{\text{Pt allele (peak height)} + \text{Donor allele (peak height)}}$$

$$\text{Percent Engraftment} = \frac{p119}{p119 + d127}$$

Blue	D8S1179	144.07	144.07		
	D21S11	208.08	217.95	I	p208.08 / (p208.08 + d204.02)
	D7S820	270.99	283.14		
	CSF1PO(5q)	325.13	325.13		
Green	D3S1358	123.63	131.89		
	TH01(11p)	185.93	185.93	I	p185.93 / (p185.93 + d178.79 + d182.78)
	D13S317	220.5	240.36	I	(p220.50 + p240.36) / (p220.50 + p240.36 + d228.06 + d232.09)
	D16S539	268.23	276.14	I	(p268.23 * 2) / ((p268.23 * 2) + d280.03)
	D2S1338	327.53	343.62	I	p327.53 / (p327.53 + d335.16)
Yellow	D19S433	117.33	129.35	I	p129.35 / (p129.35 + d121.20)
	VWA(12p)	170.34	178.45		
	TPOX(2p)	241.75	241.75	I	p241.75 / (p241.75 + d229.67 + d233.61)
	D18S51	282.18	298.54	I	p298.54 / (p298.54 + d306.72)
Red	Amelogenin	106.42	112.04		
	D5S818	155.31	155.31		
	FGA(4q)	218.32	236.38	I	p236.38 / (p236.38 + d238.11)

Blue	D8S1179	135.05	143.8	I	(d135.05 * 2) / (d135.05 + pd144.07)
	D21S11	204.02	217.83		
	D7S820	270.82	278.91		
	CSF1PO(5q)	328.65	332.73	I	(d328.65 + d332.73) / (d328.65 + d332.73 + p325.13)
Green	D3S1358	123.22	131.5		
	TH01(11p)	178.79	182.78		
	D13S317	228.06	232.09	I	(d228.06 + d232.09) / (d228.06 + d232.09 + p220.50 + p240.36)
	D16S539	280.03	280.03	I	d280.03 / (d280.03 + p268.23 + p276.14)
	D2S1338	335.16	346.95	I	(d335.16 + d346.95) / (d335.16 + d346.95 + p327.53 + p343.62)
Yellow	D19S433	113.31	121.2	I	d121.20 / (d121.20 + p129.35)
	VWA(12p)	174.27	178.27		
	TPOX(2p)	229.67	233.61	I	(d229.67 * 2) / ((d229.67 * 2) + p241.75)
	D18S51	281.9	306.72	I	d306.72 / (d306.72 + p298.54)
Red	Amelogenin	106.15	111.83		
	D5S818	151	155.13		
	FGA(4q)	234.06	238.11	I	d238.11 / (d238.11 + p236.38)

Post-Transplant

- ▶ Height values (capillary electrophoresis peaks) of donor and recipient ‘informative’ alleles are looked for in a patient post-transplant specimen to determine the ratio of donor-derived alleles to patient-derived alleles.
- ▶ Percent engraftment is manually calculated

- ▶ Problem: This manual analysis is extremely detailed and time-consuming, lending itself to human error.
- ▶ Approach: We created a computer program named *ChimerAnalyzer* to assist with engraftment chimerism analysis.

Advanced Search



ABOUT >>

NEWS >>

DOCUMENTATION >>

DOWNLOAD >>

NEWS >>

COMMUNITY >>

FOUNDATION >>

CORE DEVELOPMENT >>

Help[Package Index](#)[Quick Links \(2.7.3\)](#)

- » Documentation
- » Windows Installer
- » Source Distribution

[Quick Links \(3.3.0\)](#)

- » Documentation
- » Windows Installer
- » Source Distribution

[Python Jobs](#)[Python Merchandise](#)[Python Wiki](#)[Python Insider Blog](#)[Python 2 or 3?](#)[Help Fund Python](#)[Non-English Resources](#)**Release Schedule**[Saturday, April 6](#)[Python 2.7.4](#)

Python Programming Language – Official Website

Python is a programming language that lets you work more quickly and integrate your systems more effectively. You can learn to use Python and see almost immediate gains in productivity and lower maintenance costs.

Python runs on Windows, Linux/Unix, Mac OS X, and has been ported to the Java and .NET virtual machines.

Python is free to use, even for commercial products, because of its OSI-approved [open source license](#).

New to Python or choosing between Python 2 and Python 3? Read [Python 2 or Python 3](#).

The [Python Software Foundation](#) holds the intellectual property rights behind Python, underwrites the [PyCon conference](#), and funds many other projects in the Python community.

[Read more, -or- download Python now](#)

» Release candidates for Python 3.2.4 and 3.3.1 have been released

Release candidates for [Python 3.2.4](#) and [Python 3.3.1](#) have been released. Please test!

Published: Mon, 25 March 2013, 11:00 +0200

» A release candidate for Python 2.7.4 has been released

A [Release candidate for Python 2.7.4](#) has been released. Please test!

Published: Sat, 23 March 2013, 11:00 -0500

» Python trademark dispute resolved!

The Python Software Foundation thanks the community for their support during our European Union trademark dispute. We have settled the dispute and have now obtained the trademark within the EU, and our opponent has ceased using the Python name for their services. More details are available on the [PSF blog](#).

Published: Fri, 22 March 2013, 21:00 +0200

» Python trademark at risk in Europe: We need your help!

The Python Software Foundation is requesting the assistance of anyone in the community who works in a company that has a presence in an EU Community Member State. The Python trademark is at risk, more details are available on the [PSF blog](#). Thanks.

Published: Thu, 14 February 2013, 21:00 +0200

» Python wiki server online again after recovery

Support the Python Community

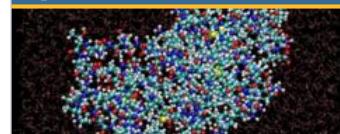
Help the Python community by becoming an associate member or making a one-time donation.

Python 3 Poll

I wish there was Python 3 support in

(enter PyPI package name)

[Vote](#) [Results](#)

Simulate biomolecules with Python...

... joining users such as Rackspace, Industrial Light and Magic, AstraZeneca, Honeywell, and many others.

What they are saying...**ITA Software:**

We knew we needed a software load balancer to meet a customer commitment to have their site up and running.

We had this hacked-together version that somebody had

DEMO

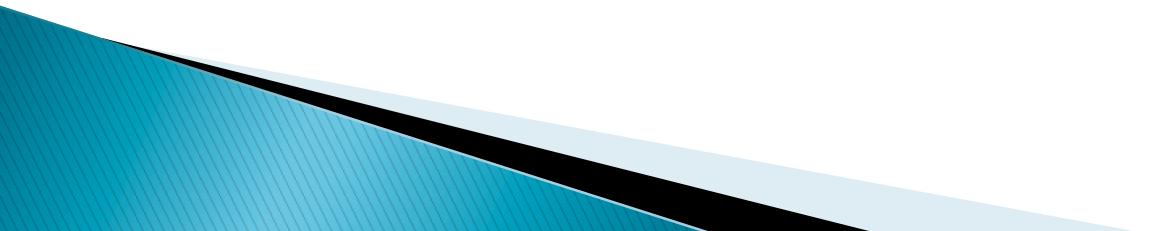


Table 2: Time Estimate for Analysis of a Single Complete Case^c

	Total Time/Case
Manual Evaluation	30 minutes
ChimerAnalyzer	2 - 2.5 minutes
Time Saved with ChimerAnalyzer	27.5 minutes (92%)

^cIncluding a pre-transplant analysis of donor and patient samples and 3 post-transplant follow-up samples.

Validation of 469 samples was performed and compared to the existing manual analysis

Table 4: “Miscall” Rate Summary: Accuracy of *ChimerAnalyzer* and Manual Method Relative to Gold Standard^c

	Total	Correctly Determined by <i>ChimerAnalyzer</i>	Correctly Determined by Manual Method
PRE-TRANSPLANT			
Alleles			7102/7104 (99.97%) ^e
Without locus exclusion tool	7104 ^d	7099/7104 (99.93%)	
With locus exclusion tool	7094	7094/7094 (100%)	
Informative Loci/Equations			
Without locus exclusion tool	1350	1346/1350 (99.70%)	
With locus exclusion tool	1346	1346/1346 (100%)	
Non-informative Loci			
Without locus exclusion tool	2202	2200/2202 (99.91%)	
With locus exclusion tool	2200	2200/2200 (100%)	
POST-TRANSPLANT			
Follow-up Calculations			
Without locus exclusion tool	1375	1365/1375 (99.27%)	
With locus exclusion tool	1365	1365/1365 (100%)	

^cThe Gold Standard represents a combination of automated and manual methods.

^dTotal theoretical alleles in 111 pre-transplant workups.

^eTwo alleles missed by the manual method and caught by *ChimerAnalyzer*.

^fCalculation errors during the manual evaluation caught by *ChimerAnalyzer*.

Conclusions

- ▶ In summary, *ChimerAnalyzer* can accurately and reliably automate bone marrow engraftment analysis, dramatically decreasing the amount of time required for an analysis.
- ▶ *ChimerAnalyzer* has the potential to achieve 100% accuracy and can analyze all informative alleles.
- ▶ *ChimerAnalyzer* is currently the method of choice for BME analysis at Hopkins and will be made freely available for public use via internet download.

Future Endeavors

- ▶ In order to prevent miscalling of true and false peaks in the future, we hope to develop a logistic regression machine learning algorithm to differentiate true peaks from false peaks in the pre-transplant analysis.
- ▶ This will allow *ChimerAnalyzer* to learn the characteristics of a true peak versus a false peak and eventually predict whether a peak is real or not based on a learned relationship as represented by a function.
- ▶ This regression algorithm will complement our current rule-based approach.

Thank You

- ▶ Dr. Christopher Gocke
- ▶ Molecular Diagnostics Lab at Johns Hopkins