

From CPI to Fully Continuous: The USACIL Journey to Better Mixture Interpretation

Creating a Quantitative Drop Model LR

Tim Kalafut, USACIL

timothy.s.kalafut.civ@mail.mil

Disclaimer

- The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.
- The US Army has licensed the ArmedXpert software package to NicheVision, LLC. The presenter receives no compensation other than normal salary from the Army for regular duties.

Where we were... CPI

- For years we just did the ol' "Sum 'em and square 'em" statistical approach using all detected alleles (pre 2010 guidelines)
- It worked well when all alleles of persons of interest (POI) in the case were found in the data
- "At least X contributors..."
- We had RMP for "simple" mixtures

CPI = RIP

- SWGDAM Guidelines came out...
- We felt like we had the rug pulled out from under us
- But wait....
 - We also used RMP
 - We had an in-house Excel program that was pretty good using RMP for 2 person mixtures
 - Also deconvoluted 2 and 3 person mixtures
 - Based on 3 simple rules
 - Could automatically condition results on a reference

RMP

- Our existing software was easily extended to 3 person mixtures
- All combinations possible at a locus were determined
- All peak height ratios and proportions were calculated
- Only the ones that make sense needed to be considered
- A good, solid quantitative binary model for deconvolution

RMP

- How do we get RMP to work for 3 people?
- We found everything we needed in the SWGDAM guidelines to use various “flavors” of RMP stats
- We extended the RMP to three contributors including dealing with drop out situations via the 2P concept – (NicheVision involved)
- Works great when interpretable loci end up including the references in the case

RMP

- We got really good at using RMP and interpreting partial/degraded/complex mixtures
- “Turbo” RMP stats served us well for 95%+ of our casework samples
- But some samples just didn’t fit...
 - Would be labeled inconclusive...
 - Then you check against POI and POI alleles present... but sample not interpretable –
FRUSTRATING!

LR to the rescue?

- We didn't use LR, even though in-house software could do it
- In the LR approach, you consider POI profiles during the interpretation process
- But how to deal with missing alleles...
 - Is there a “2p” version of the LR?
 - If we find one, can we salvage some of those samples the RMP can't handle?

LR to the rescue?

- We found out there were various LR models
- Binary models
 - UC model (becomes the mUC with drop out)
 - R model (restricted –quantitative)
- Continuous LR models
 - F model
 - Q model

What we did...

- We went looking for help...
- We got in contact with John Buckleton and Jo Bright at ESR in Auckland, NZ
- In November of 2011 we sent an email...
- Dear Dr. Buckleton,
 - I'd like to ask you a question about using the LR when drop out is a concern. We have tweaked the RMP stats to handle this, but have reached the limits of that for 3 persons using PHr and P.....

How we got started

- They invited us to their lab to teach us the Q model – visited in Feb 2012
- Kelly et al paper just published out of their lab
- We had a pretty good software programmer we'd been working with....

Q model LR

Forensic Science International: Genetics 6 (2012) 191–197



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig



The interpretation of low level DNA mixtures

Hannah Kelly^{a,*}, Jo-Anne Bright^a, James Curran^b, John Buckleton^a

^a ESR, PB 92021 Auckland, New Zealand

^b Department of Statistics, University of Auckland, PB 92019 Auckland, New Zealand

$$\begin{aligned} & \frac{4!}{2!1!1!1!} \Pr(13, 13, 14, 15|X) + \frac{4!}{1!2!1!1!} \Pr(13, 14, 14, 15|X) + \\ & \frac{4!}{1!1!1!2!} \Pr(13, 14, 15, 15|X) + \frac{4!}{1!1!1!1!1!} \Pr(13, 14, 15, Q|X) = 12\Pr(13, 14, 15|X) \times \left[\begin{array}{l} \Pr(13|13, 14, 15, X) \\ + \Pr(14|13, 14, 15, X) \\ + \Pr(15|13, 14, 15, X) \\ + 2\Pr(Q|13, 14, 15, X) \end{array} \right] \end{aligned}$$

Table 3 (Continued)

Q model	$\frac{12(\theta(1-\theta)p_{28})(\theta+(1-\theta)p_{30})}{(1+\theta)(1+2\theta)}$ <p>Allelic vector (28,30)</p> <p>$\Pr(E Hp) = 1$</p> $4\Pr(28, 28, 28, 30 28, 30) + 6\Pr(28, 28, 30, 30 28, 30) + 4\Pr(28, 30, 30, 30 28, 30) + 12\Pr(28, 28, 30, Q 28, 30)$ $+ 12\Pr(28, 30, 30, Q 28, 30)$ $+ 12\Pr(28, 30, Q, Q 28, 30)$ <p>$\Pr(E Hd) = 2\Pr(28, 30 28, 30) \times \left[\begin{array}{l} 6 - 6\Pr(28 28, 28, 30, 30) - 6\Pr(30 28, 28, 30, 30) + 2\Pr(28, 28 28, 28, 30, 30) \\ + 2\Pr(30, 30 28, 28, 30, 30) \\ + 3\Pr(28, 30 28, 28, 30, 30) \end{array} \right]$ $\frac{2(\theta(1-\theta)p_{28})(\theta+(1-\theta)p_{30})}{(1+\theta)(1+2\theta)} \times$ $\left[\begin{array}{l} 6 - \frac{6(2\theta+(1-\theta)p_{28})}{(1+3\theta)} - \frac{6(2\theta+(1-\theta)p_{30})}{(1+3\theta)} + \frac{2(2\theta+(1-\theta)p_{28})(3\theta+(1-\theta)p_{28})}{(1+3\theta)(1+4\theta)} + \frac{2(2\theta+(1-\theta)p_{30})(3\theta(1-\theta)p_{30})}{(1+3\theta)(1+4\theta)} \\ + \frac{3(2\theta+(1-\theta)p_{28})(2\theta+(1-\theta)p_{30})}{(1+3\theta)(1+4\theta)} \end{array} \right]$ </p>
---------	---

What happened next – Part 1

- We showed them our software with 2 and 3 person deconvolution and “turbo” RMP stats
- While there we started discussing a “hybrid” LR model that is both continuous and quantitative
- We call it D model
 - It’s based on the way we’ve always deconvoluted profiles
 - Adds an allele specific probability of Drop out based on each questioned sample amp

What happened next – Part 2

- While there, they showed us a piece of software they were developing – STRmix (DyNAmix originally)
 - Fully continuous
 - MCMC based
 - Really impressive
 - HUGE jump from RMP world
- We left with intentions of working on D model for us and an interest in STRmix

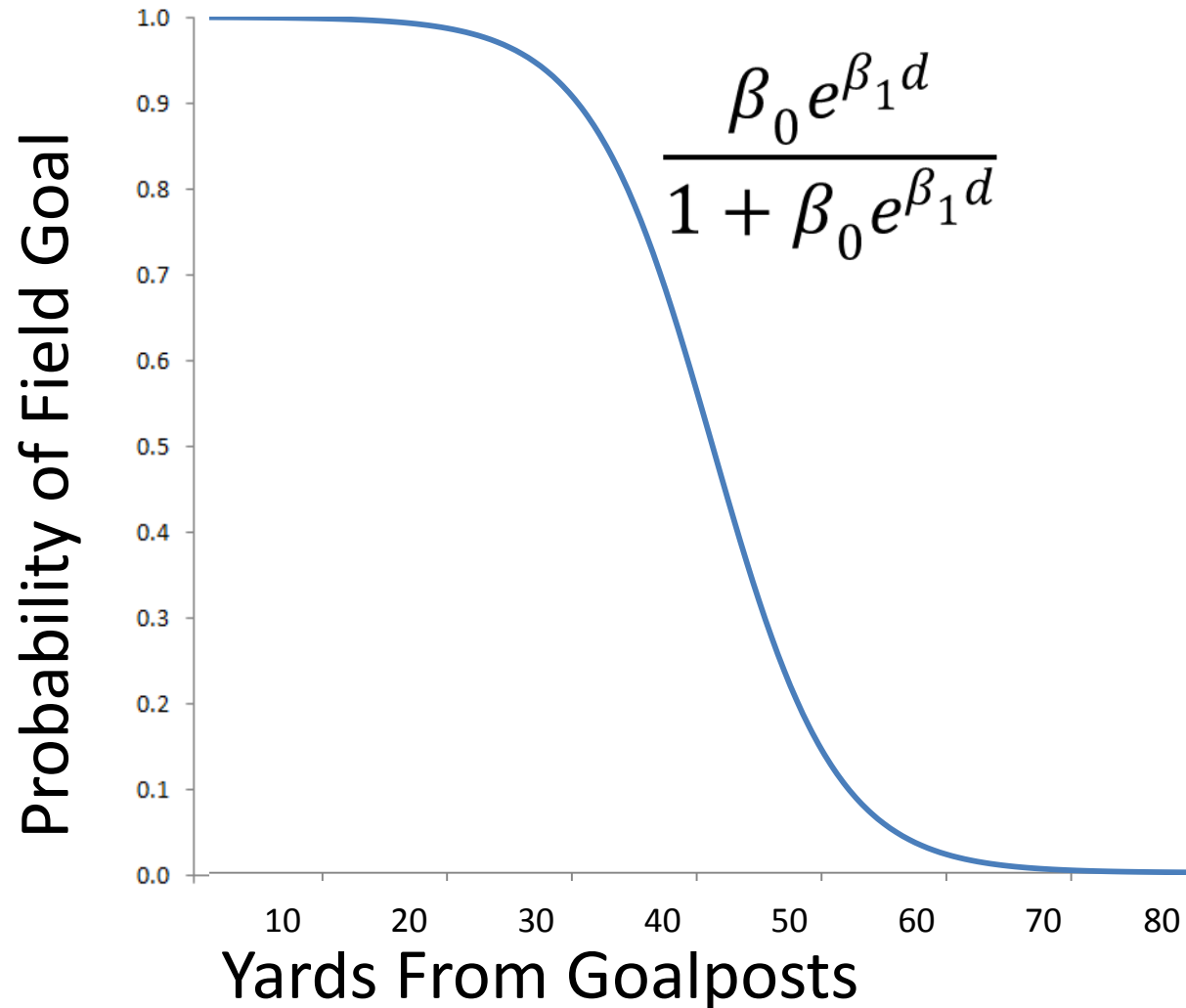
Where we are today

- STRmix was pretty much a complete product
- D model had to be developed, coded, tested, broken, re-tested, etc....
- Today at USACIL we are finishing our internal competency testing on STRmix
- We are supporting, testing, developing the ArmedXpert D Model with NicheVision and anticipate it being another tool available to use when it's fully finished

D Model Strategy

- Step 1: Validate
 - Run a bunch of samples with varying levels of drop out for which you know the true types
 - Develop a logistic regression curve that relates probability of drop out – $\Pr(D)$ – to allele height
- Step 2: Solve degradation curve of Q sample
 - Contributor specific
 - Results in allele specific probability of drop out
 - Apply quantitative information (deconvolute)
 - Build the LR

Logistic Regression in Football



Logistic Regression in Allelic Drop Out

Forensic Science International: Genetics 9 (2014) 9–11



Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig



Utilising allelic dropout probabilities estimated by logistic regression in casework



John Buckleton^{a,*}, Hannah Kelly^b, Jo-Anne Bright^{a,b}, Duncan Taylor^c,
Torben Tvedebrink^{d,e}, James M. Curran^b

^a ESR Ltd, Private Bag 92021, Auckland, New Zealand

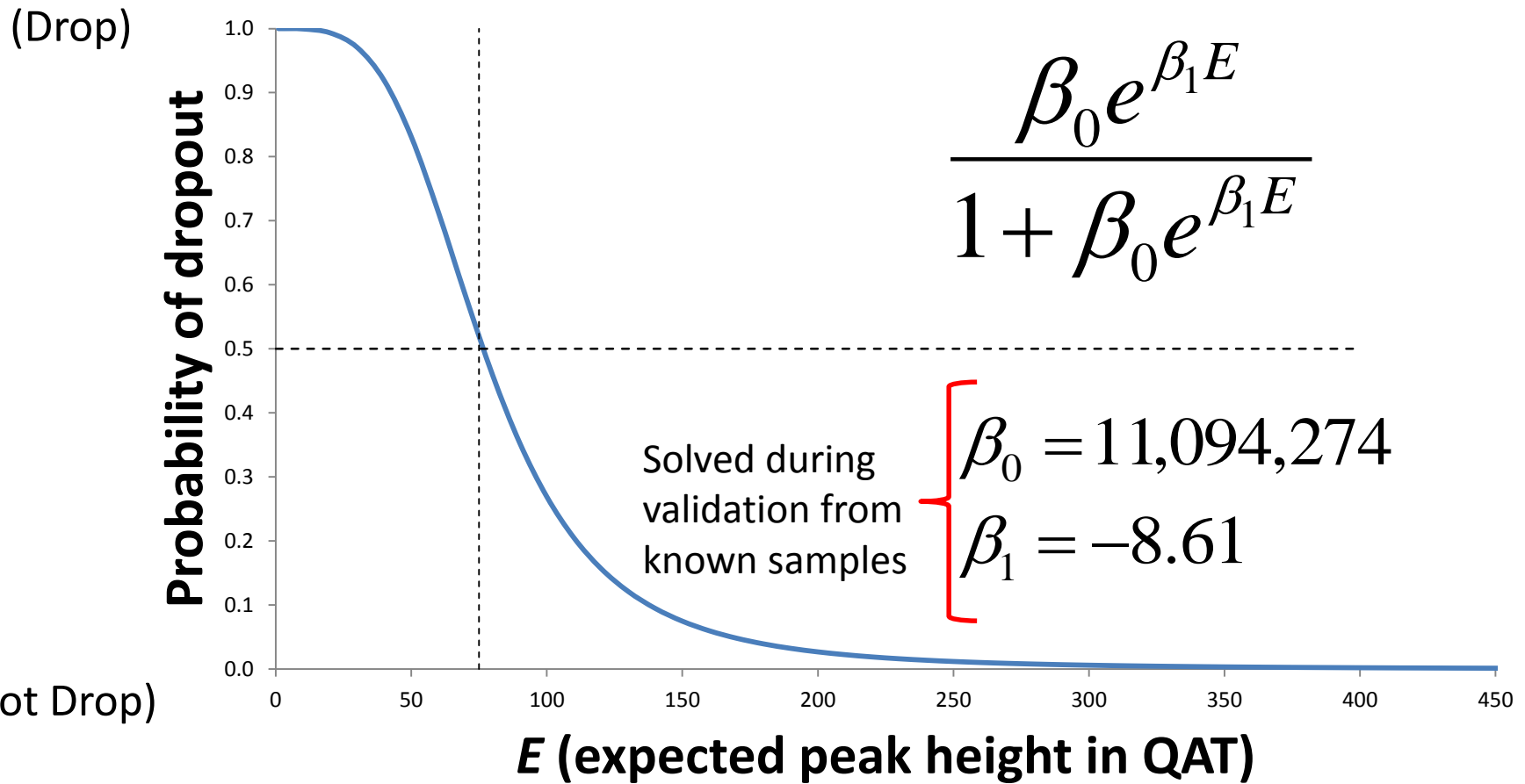
^b Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand

^c Forensic Science South Australia, 21 Divett Place, SA 5000, Australia

^d Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg East, Denmark

^e Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health Sciences, University of Copenhagen, Denmark

Results of 140 samples (75 rfu threshold)



QAT

- Q: What determines the $\text{Pr}(D)$?
- A: The amount of template available for the enzyme to amp
- Note this is not part of the quant step, and the true value both varies across the profile (degradation) and can never be known
- But we can make a proxy by plotting a curve based on the observed rfu height
- Results in “Quality Amplifiable Template”

Degradation Curve

Australian Journal of Forensic Sciences, 2013
<http://dx.doi.org/10.1080/00450618.2013.772235>



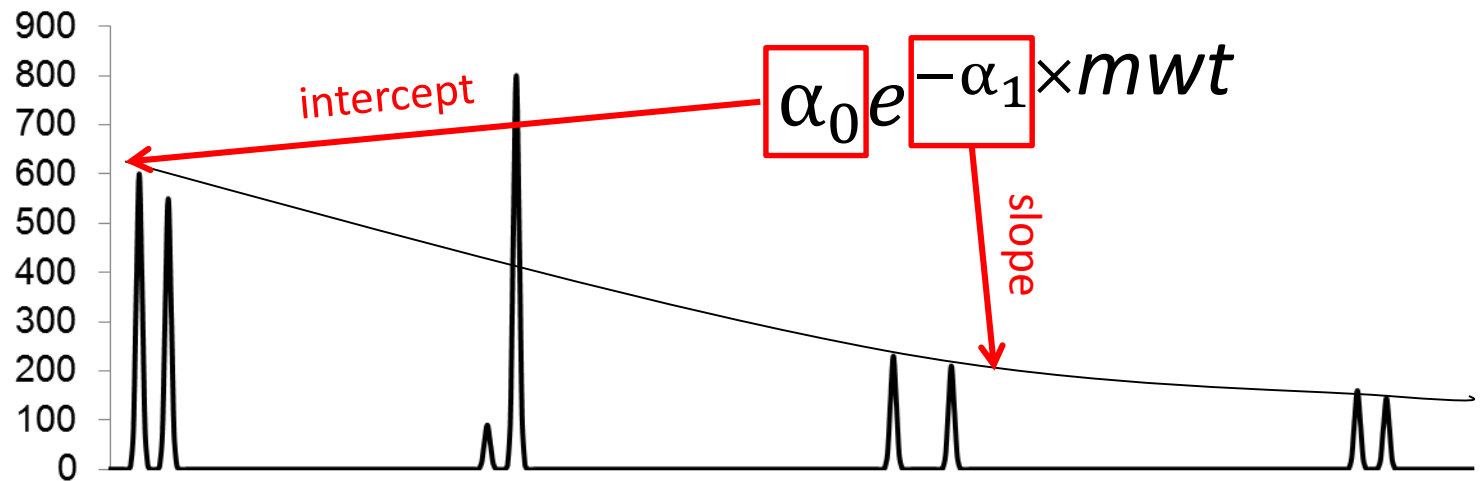
Degradation of forensic DNA profiles

Jo-Anne Bright^{a,b*}, Duncan Taylor^c, James M. Curran^b and John S. Buckleton^a

^aESR, Private Bag 92021, Auckland 1025, New Zealand; ^bDepartment of Statistics, University of Auckland, Private Bag 92019, Auckland 1025, New Zealand; ^cForensic Science South Australia, 21 Divett Place, SA 5000, Australia

Degradation curve

- Empirical data has shown that for larger multiplexes a DNA slope is best described by an *exponential curve*



QAT related to RFU $E = \alpha_0 e^{-\alpha_1 \times mwt}$

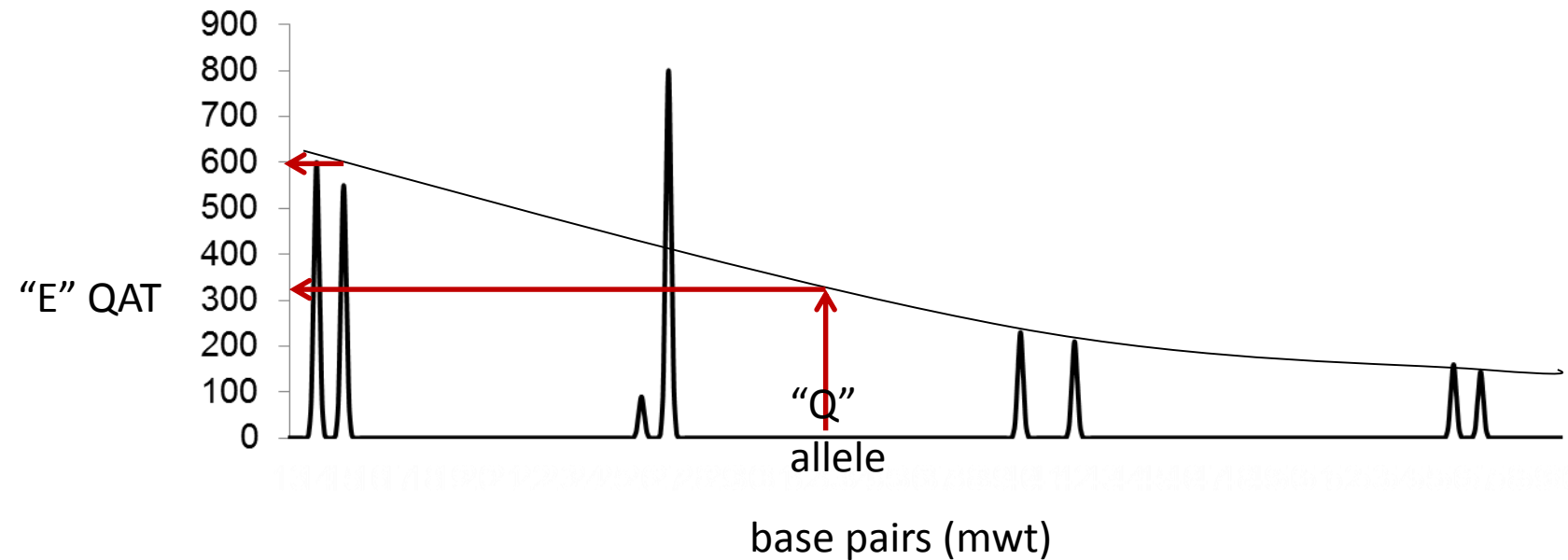
- The “*mwt*” term in the equation corrects for allele size (bp)
- The α values are solved by the software
- Detected peaks:
 - QAT could be higher or lower than the rfu value
 - It’s possible for all peaks to have the same QAT value across the profile

QAT related to RFU

- Dropped peaks:
 - New term to get used to: EXPECTED PEAK HEIGHT or how tall should that peak have been
 - In other words, if we know the bp size of the allele that dropped, we can determine how tall (in QAT) it should have been
- (RFU is *observed* peak height, while QAT is *expected* peak height)

Degradation curve

- Once you have the curve, you can now determine E (Expected height in QATs) for any allele



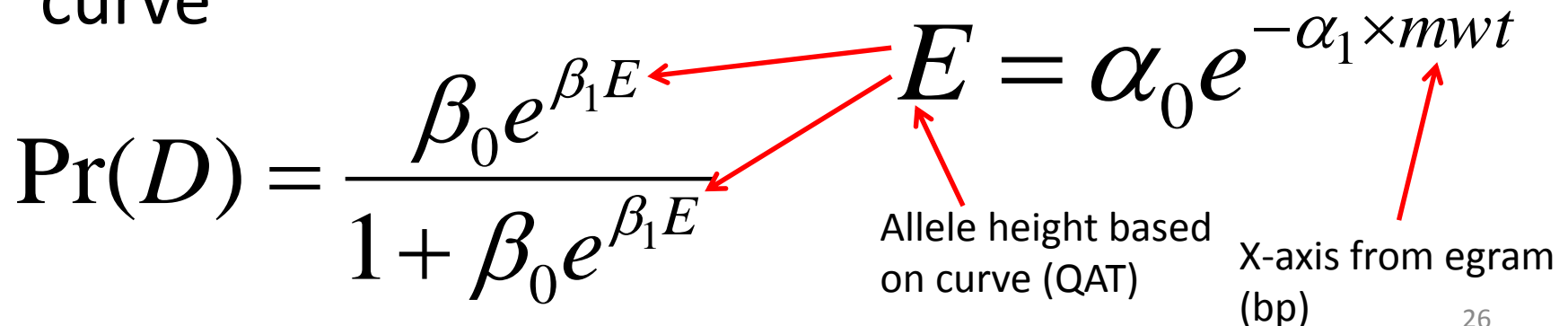
Relating back to the Pr(D) curve

- The logistic regression curve has an E in it
 - Expected peak height from the degradation curve of the particular sample in question
 - This E and the values for β from validation give Pr(D)
- Remember, the E comes from the degradation curve

$$\Pr(D) = \frac{\beta_0 e^{\beta_1 E}}{1 + \beta_0 e^{\beta_1 E}}$$

$E = \alpha_0 e^{-\alpha_1 \times mwt}$

Allele height based on curve (QAT) X-axis from egram (bp)

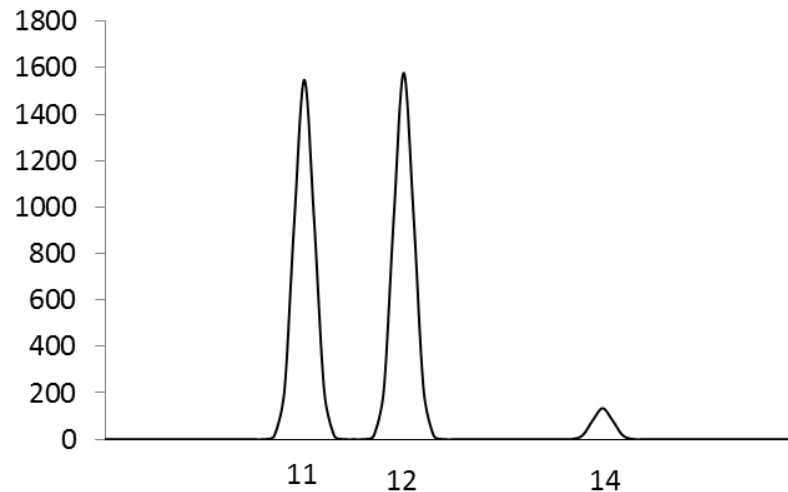


Locus example – one drop

D13S317

POI = 13,14

$$f_{11} = 0.1, f_{12} = 0.1, f_{14} = 0.1, f_Q = 0.7$$



β_0	β_1
11,094,274	-8.61

E_{13}	E_{14}	E_Q
104	103	105

$\Pr(D_{13})$	$\Pr(D_{14})$	$\Pr(D_{14,14})$	$\Pr(D_Q)$
0.242	0.251	0.025	0.238

Deconvolute

- This is the ArmedXpert deconvolution window
- It's set for 2 people
 - Limited to 50% phr
 - Conditioned on V

Mixture Interpretation - DDA Interpretation

Setup

Pick via mouse Vaginal swab Operations Contributor # 2

References

Locus D13S317 (3) <- -> Vaginal swab 05 n/a Highest to lowest # PHr 0.50

Alleles 11, 12, 14 Vaginal swab 06 n/a mPH 75

RFUs 1591, 1622, 134 Ignore alleles below mPH mP 0.00

BPs 228, 232, 240 Look locus on report Popout calls View call report

Add Comment

Apply Globally Apply Stutter Add Profile 0.058

11 12 14
1591 1622 134

Mixture Information 12

Only combinations including the following reference profiles are included: (11, 12)

All combinations have: PHr >= 0.5, MPh >= 75, mP >= 0

For a 2-contributor 3-allele mixture of types AB & AC: 2/3-combination(s):

11, 12(phr = 0.91; p = 0.92) [Ref. 1] • 11, 14(phr = 0.91; p = 0.08) [11.5 : 1]

11, 12(phr = 0.94; p = 0.92) [Ref. 1] • 12, 14(phr = 0.94; p = 0.08) [11.5 : 1]

For a 2-contributor 3-allele mixture of types AB & CC: 1/3-combination(s):

11, 12(phr = 0.98; p = 0.96) [Ref. 1] • 14(p = 0.04) [24 : 1]

Deconvolute

- Although there are really 6 different ways two people can make a three allele pattern, only 3 fit our constraints

- In this case, that's true even without conditioning on a V profile

- The “Q” or drop allele isn't shown - yet

Mixture Interpretation - DDA Interpretation

Setup

Pick via mouse Vaginal swab

Operations Contributor # 2

References

Locus D13S317 (3) <- ->

Vaginal swab 05 n/a

Vaginal swab 06 n/a

Highest to lowest #

Ignore alleles below mPH

Look locus on report

PHr 0.50

mPH 75

mP 0.00

HT 300

Popout calls

View call report

Add Comment

Apply Globally

Apply Stutter

Add Profile

0.058

11 12 14

1591 1622 134

Mixture Information

Only combinations including the following reference profiles are included: (11, 12)

All combinations have: PHr >= 0.5, mPH >= 75, mP >= 0

For a 2-contributor 3-allele mixture of types AB & AC: 2/3-combination(s):

11, 12(phr = 0.91; p = 0.92) [Ref. 1] • 11, 14(phr = 0.91; p = 0.08) [11.5 : 1]

11, 12(phr = 0.94; p = 0.92) [Ref. 1] • 12, 14(phr = 0.94; p = 0.08) [11.5 : 1]

For a 2-contributor 3-allele mixture of types AB & CC: 1/3-combination(s):

11, 12(phr = 0.98; p = 0.96) [Ref. 1] • 14(p = 0.04) [24 : 1]

Make a list of genotypes you care about

- These are the only genotypes the minor foreign contributor could be based on the settings we told AX to use
- Note $p = 0.08$ or 0.04 , minor proportion is 8% or 4% (not counting potential drop or 14,Q)

1] • 11, 14(phr = 0.91; p = 0.08) |
1] • 12, 14(phr = 0.94; p = 0.08) |
of types AB & CC: 1/3-combinatio
1] • 14(p = 0.04) [24 : 1]

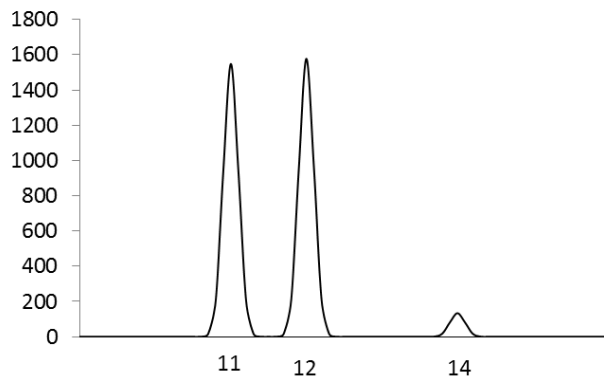
Remember the LR

- 2 competing propositions
 - H_1 or H_p = What prosecution thinks the evidence explains
 - H_2 or H_d = What defense thinks
- $LR > 1$ in favor of prosecutor/numerator
- $LR < 1$ in favor of defense/denominator

$$LR = \frac{\Pr(E \mid H_1, I)}{\Pr(E \mid H_2, I)}$$

What we've done so far

- We've determined a degradation curve for this sample
- That degradation curve gave us our expected peak heights for both detected and any dropped alleles
- We've then compared that *Expected* height to the *Beta* curve (log regression) to determine $Pr(D)$ or $Pr(N)$
- Now build the LR – I'll start with defense

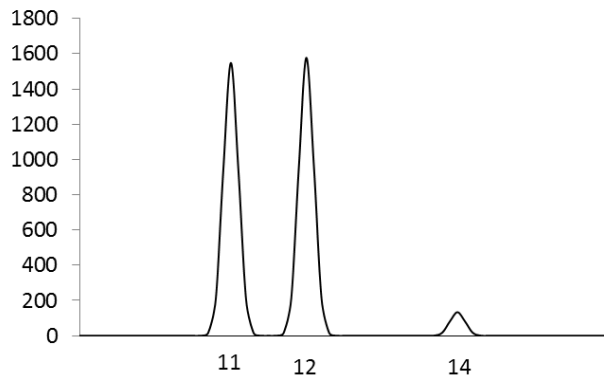


D13S317

H₂ - Defense

List of Genotypes				
14,14				
11,14				
12,14				
14,Q				

Make a list of Genotypes you care about.....



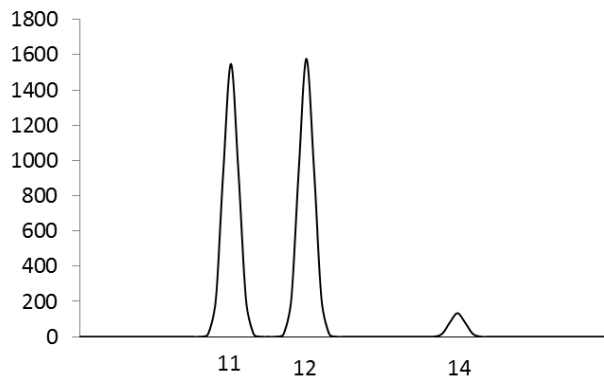
D13S317

H₂ - Defense

f_{11}	f_{12}	f_{13}	f_Q
0.10	0.10	0.10	0.70

List of Genotypes	Genotype Frequency			
14,14	0.01			
11,14	0.02			
12,14	0.02			
14,Q	0.14			

.....and calculate those genotype frequencies



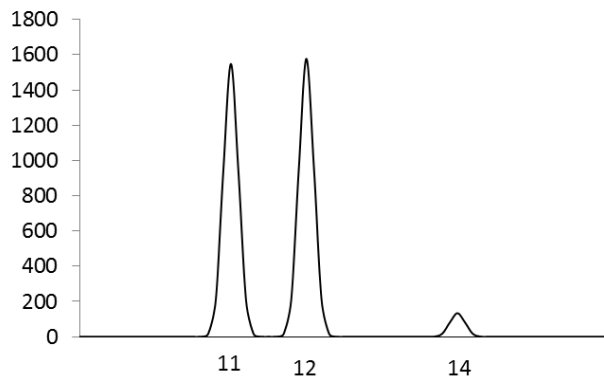
D13S317

H₂ - Defense

Pr(D ₁₄)	Pr(D _{14,14})	Pr(D _Q)
0.251	0.025	0.238

List of Genotypes	Genotype Frequency	Drop or Not Drop	Modifying value based on Pr(D)	
14,14	0.01	$\bar{D}_{14,14}$	1-0.025	
11,14	0.02	\bar{D}_{14}	1-0.251	
12,14	0.02	\bar{D}_{14}	1-0.251	
14,Q	0.14	$\bar{D}_{14}D_Q$	(1-0.251) x 0.238	

Modify genotype freqs by Pr(D) and/or Pr(N) as needed



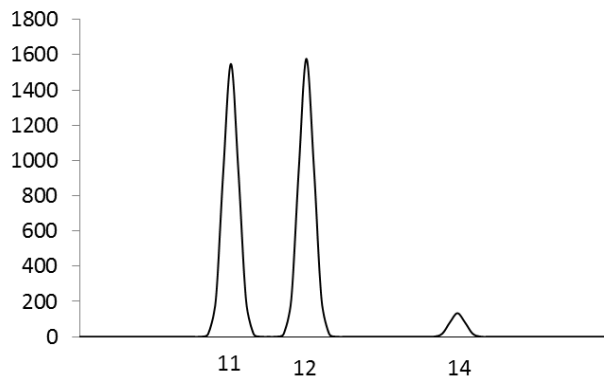
D13S317

H₂ - Defense

Pr(D ₁₄)	Pr(D _{14,14})	Pr(D _Q)
0.251	0.025	0.238

List of Genotypes	Genotype Frequency	Drop or Not Drop	Modifying value based on Pr(D)	Multiply Across
14,14	0.01	$\overline{D}_{14,14}$	1-0.025	0.00975
11,14	0.02	\overline{D}_{14}	1-0.251	0.0150
12,14	0.02	\overline{D}_{14}	1-0.251	0.0150
14,Q	0.14	$\overline{D}_{14}D_Q$	(1-0.251) x 0.238	0.0250

Multiply across the rows.... (2pq x Pr(D))



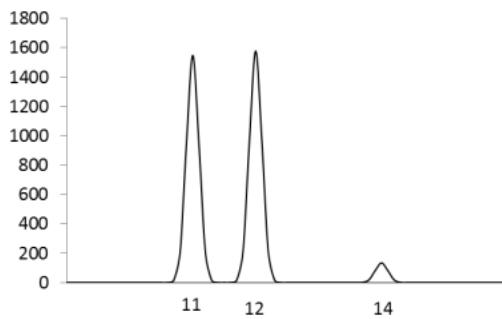
D13S317

H₂ - Defense

Pr(D ₁₄)	Pr(D _{14,14})	Pr(D _Q)
0.251	0.025	0.238

List of Genotypes	Genotype Frequency	Drop or Not Drop	Modifying value based on Pr(D)	Multiply Across
14,14	0.01	$\overline{D}_{14,14}$	1-0.025	0.00975
11,14	0.02	\overline{D}_{14}	1-0.251	0.0150
12,14	0.02	\overline{D}_{14}	1-0.251	0.0150
14,Q	0.14	$\overline{D}_{14}D_Q$	(1-0.251) x 0.238	0.0250
			Add Down:	0.06475

.....Add down to get the H₂ value for the locus



D13S317 LR
 H_1 = Prosecution

$\Pr(D_{13})$	$\Pr(D_{14})$
0.242	0.251

List of Genotypes	Genotype Frequency	Drop or Not Drop	Modifying value based on $\Pr(D)$	Multiply Across
13,14	1	$D_{13}\bar{D}_{14}$	$0.242 \times (1-0.251)$	0.181

-Note this significant difference from H_2 : The genotype probability is 1. This is because the prosecution is 100% certain the POI is the suspect in the case. (Otherwise, why are we at court in the first place?)

-However, because the 13 allele has dropped, the H_1 is penalized By the probability of drop, and the overall H_1 value is no longer 1.

-The magic in a probabilistic LR happens in the numerator!!!

LR for D13

- Take H_2 from 2 slides prior
- Divide by H_1 from previous slide

$$LR = \frac{0.181}{0.06475} = 2.795$$

- FYI – 14, Any (2p) for this locus is 14, so an LR of ~ 3 is a significant penalty to H_p compared to the RMP

D Model Summary

- Step 1: Validate your $\Pr(D)$ using logistic regression to generate your beta curve (one time)
- Step 2: Use it on a sample
 - 2A: Hang a degradation curve (alpha) on a sample to convert to QAT and find $\Pr(D)$ from beta curve
 - 2B: Deconvolute to eliminate silly combinations
 - 2C: Use that (partial) deconvolution to make a list of genotypes you care about and find those frequencies
 - 2D: Modify by $\Pr(D)$ or $1-\Pr(D)$ as needed per allele
 - 2E: Multiply across and add down

Current output example

- A summary page gives the locus by locus LR, total LR, info about set up and average mwt for each locus (used for Q allele)

Sample Type	Sample Name	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539
Mixture	Case 1 Q2	13,13	28,28	10,10	10,10	14,18	6,9,3	9,11	11,11
Hp Profile 1	Case 1 Known	13,14	28,30	10,11	10,11	14,18	6,9,3	9,11	11,11
FBI Hispanic	LR = 243647601704.0	3.255587176	8.162897975	3.451967172	3.558692136	75.61665381	8.91657227	11.29218761	9.609442191
FBI Black	LR = 4.50834875948e+13	3.512822358	3.614927766	3.488889003	3.525532985	75.15722892	43.57070654	75.4892458	10.92415947
FBI Caucasian	LR = 138518493939.0	3.280479898	4.156136365	3.440974576	3.558424862	21.90192842	7.22505348	20.49528916	12.65043335
Sample Type	Sample Name								
Mixture	Case 1 Q2								
Hp Profile 1	Case 1 Known								
No. of HP Contributors	1								
No. of HD Contributors	1								
Min PHR	0								
Min RFU	50								
Alpha Calc. PHR Ratio Filter	1								
Colors	Blue, Green, Yellow, Red								
Average Allele MWT for Ladder		146.22	212.36	273.21	323.33	125.73	180.39	230.53	275.33

Current output example

- A rather busy looking page summarizes observed and expected peak heights for detected alleles and Q alleles, Pr(D) for homs and hets and off this screen shot are the alpha values for this sample

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	
1	Profile	Marker	Allele1	Allele2	O1	O2	mwt1	mwt2	mwt1(dropped)	mwt2(dropped)	mwtQ	E1	E2	E1(dropped)	E2(dropped)	Eq	Eqq	Pr(Da1)	Pr(Da2)	Pr(Daa1)	Pr(Daa2)	
2		D8S1179	13	Q	82	0	146.9			151.1	146.223333	49.1720419	0	0	31.3592177	24.614811	49.229622	0.83973338	NA	0.28178539	NA	0.
3		D21S11	28	Q	92	0	201.9			210.02	212.362917	44.8933464	0	0	31.3592177	22.0610767	44.1221533	0.8804452	NA	0.35543909	NA	0.
4		D7S820	10	Q	83	0	273			276.98	273.212	39.9059164	0	0	31.3592177	19.9462844	39.8925688	0.91960621	NA	0.46136228	NA	0.
5		CSF1PO	10	Q	116	0	320.9			325.03	323.332	36.8613826	0	0	31.3592177	18.3575207	36.7150414	0.93898948	NA	0.53541305	NA	0.
6		D3S1358	14	18	124	138	121	137			125.7275	184.222764	178.298999	0	0	182.441804	364.883607	0.03617066	0.04068139	0.00280223	0.00316535	0.
7		TH01	6	9.3	229	201	172	187			180.385	166.039535	161.04465	0	0	163.213438	326.426876	0.05244692	0.05842182	0.00412749	0.00462457	0.
8		D13S317	9	11	217	197	209	217			230.52875	153.978466	151.488789	0	0	147.36076	294.721521	0.06836432	0.07235018	0.00546474	0.00580621	0.
9		D16S539	11		212		276.1				275.326667	268.613856	268.613856	0	0	134.505028	269.010056	0.00907733	0.00907733	0.00068547	0.00068547	0.

Current output example

- Each locus gets it's own summary page for each population group
- This is single source example
- Only a 13 was detected, but POI is 13,14

1	locus	D8S1179				Hp Profile 1			
2	Mixture	13,13			Allele	13	Q	QQ	
3	Hp Profile 1	13,14			Var	a			
4					PrDrop	0.281785	0.028101	0.971899	
5					PrNotD	0.718215	0.971899	0.028101	
6	Hp Calculations				Freq	0.3251	0.3498	0	
7	Mixture	a,a							
8	Hp Profile 1	a,b							
9									
10	HH = Mixture+Hp Profile 1								
11	Hp Profile 1	Equation	Number Subs						
12	Na,Db	(P(Na)*P((0.718214606848*0.971898797064)							
13									
14	Hp = [(P(Na)*P(Db))]								
15	Hp = [(0.718214606848*0.971898797064)]								
16	Hp = 0.698031912429								
17									
18	Hd Calculations								
19	Mixture	a,a							
20									
21	Hd = Mixture+Hd UnKnown 1								
22		Hd UnKno	Equation	Number Subs					
23	Na,Na	((F(a)^2+F(((0.3251)^2+(0.3251)*(1-0.3251)*0.01)*(0.718214606848)^2)							
24	Dq,Na	(2*F(q)*F((2*0.3498*0.3251*0.971898797064*0.718214606848)							
25									
26									
27	Hd = [((F(a)^2+F(a)*(1-F(a))*T)*P(Na)^2)+(2*F(q)*F(a)*P(Dq)*P(Na))]								
28	Hd = [(((0.3251)^2+(0.3251)*(1-0.3251)*0.01)*(0.718214606848)^2)+(2*0.3498*0.3251*0.971898797064*0.718214606848)]								
29	Hd = 0.214410450315								
30									
31	LR = Hp/Hd = 3.25558717593								
32									

Current output example

- Calculated values use in the stat are at the top
- Prosecution setup (Pr(N) for 13 x Pr(D) for 14)
- Defense setup (13,13 and 13,Q)
- LR

locus	D8S1179				Hp Profile 1		
Mixture	13,13			Allele	13 Q	QQ	
Hp Profile 1	13,14			Var	a		
				PrDrop	0.281785	0.028101	0.971899
				PrNotD	0.718215	0.971899	0.028101
				Freq	0.3251	0.3498	
Hp Calculations							
Mixture	a,a						
Hp Profile 1	a,b						
Hp = Mixture+Hp Profile 1							
Hp Profile 1	Equation	Number Subs					
Na,Db	(P(Na)*P(Db))	(0.718214606848*0.971898797064)					
Hp = [(P(Na)*P(Db))]							
Hp = [(0.718214606848*0.971898797064)]							
Hp = 0.698031912429							
Hd Calculations							
Mixture	a,a						
Hd = Mixture+Hd UnKnown 1							
Hd UnKnown 1	Equation	Number Subs					
Na,Na	((F(a)^2+F(((0.3251)^2+(0.3251)*(1-0.3251)*0.01)*(0.718214606848)^2)						
Dq,Na	(2*F(q)*F((2*0.3498*0.3251*0.971898797064*0.718214606848)						
Hd = [((F(a)^2+F(a)*(1-F(a))*T)*P(Na)^2)+(2*F(q)*F(a)*P(Dq)*P(Na))]							
Hd = [(((0.3251)^2+(0.3251)*(1-0.3251)*0.01)*(0.718214606848)^2)+(2*0.3498*0.3251*0.971898797064*0.718214606848)]							
Hd = 0.214410450315							
LR = Hp/Hd = 3.25558717593							

H_p

H_d

D Model Summary

- It is probabilistic – deals with “maybe”
 - Allele specific probability of drop per contributor per sample
- It is quantitative
 - Only considers genotype combinations that make sense
 - Can be more restrictive at high RFU and less at low
- It is fully continuous
 - Well, almost (semi-continuous isn't quantitative)
 - At some level peaks are so low you have consider all options so some thresholds on combinations

Impact of probabilistic on casework

- D model isn't in use yet although trials against STRmix look good; for now we use STRmix
- So the impact of STRmix.....
- We can use more samples
- We still interpret for inclusion/exclusion – we are the experts, not the software
- Early discussions with lawyers show they like the “X times more likely”

Unexpected side effects

- We (almost) always stated # of contributors for every sample, but now we must (no more “additional genetic data at 2 loci” and doing a Single Source stat)
- We still need to interpret first, STRmix only gives weight to what the expert interprets
- The maths have been adopted quite readily
- Determining which propositions to include in the LR is challenging

Unexpected side effects

- We expected that our existing ArmedXpert software may be diminished somewhat, but that absolutely is not the case
- Determining # of contributors and whether or not a trace level contributor in a 4 person mixture could be Suspect X requires a thorough knowledge and training in “old school” mixture interpretation
- The term “complex mixtures” is somewhat outdated as they are either interpretable or not

Unexpected side effects

- We decided we needed to really investigate our low level data and analytical thresholds (AT) from the instruments
- Resulted in normalizing our four 3130s – each one has a slightly different injection set up, but all give similar rfu and this is monitored
- We are about to go on-line with OSIRIS as our analysis software, AT is now color specific from 24-53 rfu

Thank you!

- Please feel free to contact with questions
- timothy.s.kalafut.civ@mail.mil
- 404-469-7289