AN ASSESSMENT OF LINKAGE BETWEEN FORENSIC MARKERS: CORE STRS, MINI-STRS AND INDELS.

Manuel Fondevila[1,2*], Carla Santos[1], Christopher Phillips[1], Ángel Carracedo[1], John M. Butler[2], Maria Victoria Lareu[1], Peter M. Vallone[2].

[1]*Instituto de Ciências Forenses Luis Concheiro, University of Santiago de Compostela*
[2]*U.S. National Institute of Standards and Technology, Biochemical Science Division*

*corresponding author contact e-mail: mafondevila@hotmail.com

INTRODUCTION

As the longer amplicons of conventional markers such as Short Tandem Repeats (STRs) tend to fail in the amplification of degraded DNA samples [1-2], mainstream STR typing methods yield reduced information. It is for this reason that there are several alternative strategies and markers resistant to DNA degradation that can provide information even from challenging DNA samples [3-6]. These alternative marker approaches rely on short amplicon systems to cope with the randomized postmortem DNA fragmentation and base modification as well as the inhibition of the DNA polymerases from by-products of degradation [7-9]. To overcome such effects several of the core STR markers have been redesigned to allow typing with smaller PCR fragments, the so called miniSTRs [3-4]. Several studies have been published related to the application of autosomal single nucleotide polymorphisms (SNPs) to forensic casework [1,10]. Recently, a third approach to short amplicon marker application to challenging DNA cases has been developed, known as short insertion deletion polymorphisms or InDels [6, 11, 12].

Although individually less informative than STRs [13-14], both SNPs and InDels have been shown to have greater amplification success rates than STRs from highly compromised DNA samples [1,19,20] – including both long and short amplicon STR assays. For this reason there will be cases in which a partial profile of STRs is obtained and the information would need to be supplemented by the addition of InDel and SNP genotyping. Moreover, in cases involving complex pedigrees, such as incest or differentiating relationships between close members of the same family, STR typing may require the extra power supplied by the addition of SNP [10] or InDel typing. In order to successfully combine these markers with STRs for an optimal calculation, it is important to be sure that all the markers included in the calculation are independent and therefore can be statistically combined [15].

With such a high number of markers (23 core STRs, 68 InDels in two independent assays [6,16]) incorporated in independent tests, the chances of two or more loci being located in close genomic positions are high. Keeping this in mind, we have conducted a thorough search of the genomic positions of the above forensic markers and compared these to newly published InDel loci in order to objectively assess whether they are under linkage disequilibrium or truly independent and therefore valid for statistical combination. For this goal, we made use of the whole genome SNP data from the HapMap project [17] following the guidelines recently published by Phillips *et al.* [15]. The HapMap database contains SNPs located in dense and regular distributions throughout the 22 autosomes, with their physical position, combined recombination rate in Centimorgans (cM) per megabase (Mb) for each position and the accumulated genetic distance in cM per position. We have applied the information published in the most recent HapMap recombination rates by selecting the closest SNP in this database to each of the markers of forensic interest. Using each proximal SNP we have calculated the recombination rates across the genome span between each marker pair of interest, and therefore, an estimation of their status of linkage or independence [15].

MATERIALS AND METHODS

Markers

We made use of the following 23 STR markers in common usage (herein core STRs): D1S1656, D10S1248, TH01, vWA, D12S391, D13S317, PentaE, D16S539, D18S51, D19S433, TPOX, D2S441, D2S1338, PentaD, D21S11, D22S1045, D3S1358, FGA, D5S818, CSF1PO, SE33, D7S820 and D8S1179. We also included the 23 component markers of the NIST mini-STR

assays [4,18]: D1GATA113, D1S1627, D1S1677, D10S1435, D11S4463, D12ATA63, D14S1434, D17S974, D17S1301, D18S853, D2S1776, D20S482, D20S1082, D3S4529, D3S3053, D4S2408, D4S2364, D5S2500, D6S1017, D6S474, D8S1115, D9S1122 and D9S2157.

The following InDel markers have been included, from the Qiagen commercial investigator DIPplex assay [16]:

rs2307433, rs1305047, rs2307581, rs16438, rs8177524, rs6481, rs16388, rs2307924, rs1611001, rs2067235, rs16363, rs17878444, rs2307956, rs2307959, rs28369942, rs2308292, rs1610937, rs1610905, rs1610935, rs1305056, rs2307652, rs1611048, rs17879936, rs2308072, rs3081400, rs8190570, rs17174476, rs2307570, rs17238892 and rs2308163.

Plus those of the human identification 38 InDel assay developed by R. Pereira et al [6]:

rs34541393, rs16402, rs16624, rs2307689, rs35769550, rs2307700, rs140809, rs3047269, rs33972805, rs33917182, rs1610871, rs2067238, rs2067294, rs2307710, rs2308242, rs2307580, rs1160956, rs34511541, rs2307978, rs2308137, rs35605984, rs36040336, rs1160886, rs2308026, rs2307526, rs34811743, rs2308189, rs589447, rs2308171, rs3051300, rs10629077, rs10688868, rs2067208, rs2307579, rs2308020, rs3080855, rs1610919 and rs2307839.

Information regarding these markers is displayed in Table 1. All genomic positions for the above markers were taken from the reference assembly Genome build 37.1 (GRCh37).

Statistical evaluation of linkage/independence of the markers.

As linkage between two given genomic positions is dictated by physical distance as well as recombination rate, we first surveyed the distances between marker pairs and selected only those separated by less than 10 Mb as candidates for linkage disequilibrium. We concentrated on recombination rate estimates for the resulting sub-set of loci to estimate recombination rates between them.

For each marker (InDel or STR), the closest HapMap SNP was identified and assigned as a proxy. Using the proxy positions the accumulated genetic distance in cM was determined. HapMap provides this parameter from the p-arm to the q-arm of each chromosome giving, for each SNP position, an accumulated recombination rate based on the local rate (termed the combined recombination rate) of all preceding SNPs. Thus, the difference in the accumulated recombination rate between two positions in the genome provides an accurate estimation of the linkage between those loci. A steep gradient in the cumulative rate indicates high recombination rates in that part of the chromosome, so chromosomes typically show the classic sigmoid curve of steep gradients (high recombination rates) at the telomeres and flat gradients (low recombination rates) around the centromere. In order to obtain a recombination rate: *Rc* from the map distance in cM we used the Kosambi mapping function calculator included in Phillips *et al.* [15].

RESULTS

All information about the markers employed in the analysis as well as the final statistical values for each pair is shown in Table.1.

DISCUSSION

Of a total of 114 forensic markers (STRs plus InDels) distributed through the 22 human autosomes, 32 pairs were separated by less than 10 Mb (values highlighted in bold under Physical distance in nucleotides in Table 1). Those would be considered to be at risk of linkage disequilibrium between them due to minimum physical distance. Although there is a positive relationship between distance and recombination rates between any two loci, the existence of recombination hotspots and the irregular distribution of recombination in general across the chromosomes prompted this more accurate assessment of recombination rates using Kosambi´s mapping function and high density SNP data.

From the Rc estimates made for the physically closest marker pairs, 12 of the 32 pairs were found to have such low recombination that treating them as statistically independent would not be advisable for identity and relationship testing purposes. Each of the relevant marker pairs are highlighted in gray in Table 1.

In general, low Kosambi adjusted Rc values are already considered to indicate high linkage between the two loci in question [15], a value of 0.5 represents full independence. In this work, we made special note to those marker pairs with adjusted Rc value lower than 0.04, meaning that both loci would be separated in one of every 25 recombinations. A significant loss of independence should be expected of marker pairs on such situation.

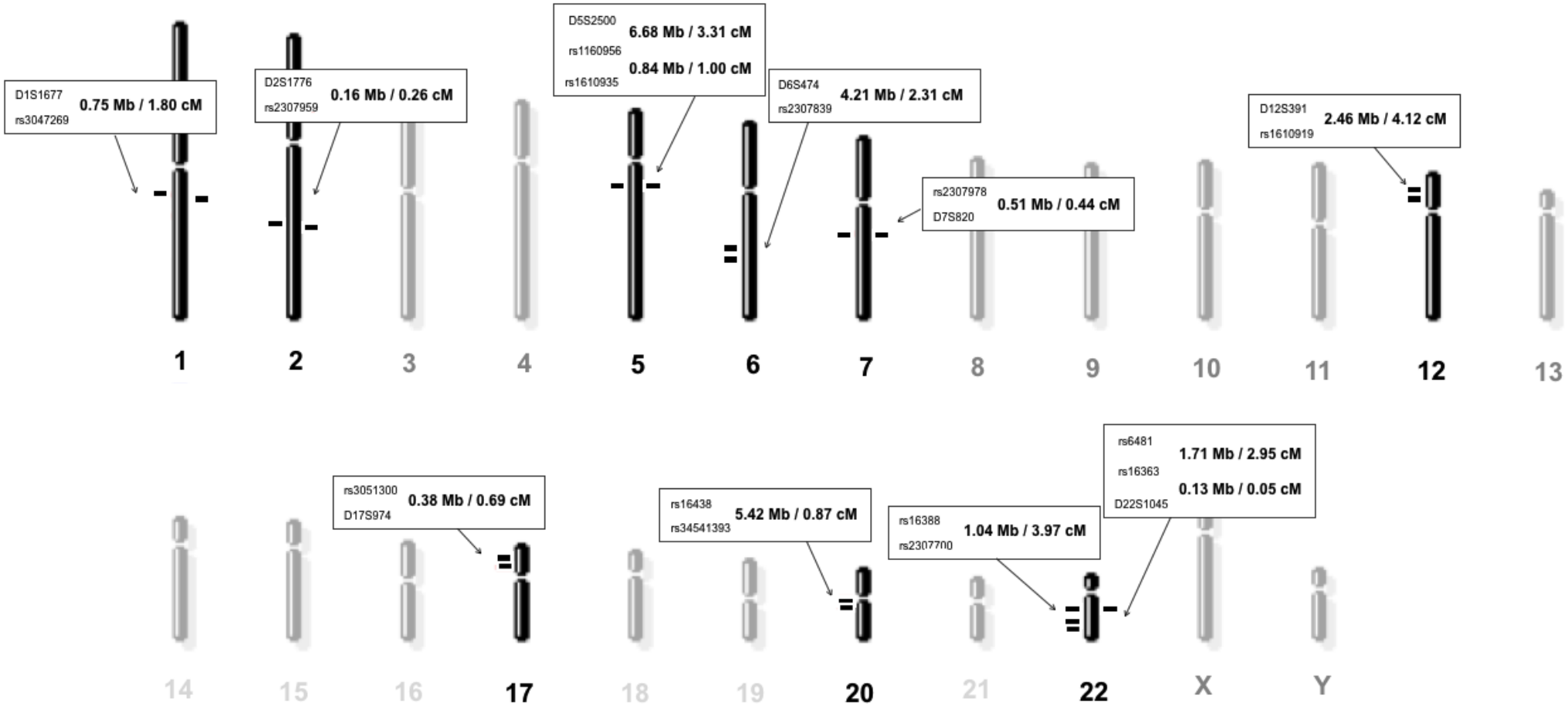

Fig.1. Chromosome distribution of markers under risk of Linkage Disequilibrium.

Table 1. Marker information and recombination mapping data.

Marker	*Chr*	*Location*	*Assay*	*class*	*SNP identifier in dbSNP or closest SNP to marker*	*HapMap proxy SNP position (bp)*	*Physical distance in nucleotides*	*cM interval of closest HapMap SNP proxies*	*Rc from Kosambi mapping function*
D1GATA113	1	7,442,859	NIST miniplex	STR	N/A	7,442,845			
rs2307956	1	54,718,192	DIPplex	InDel	rs2307956	54,719,435	47,276,590	62.4069	0.423880
rs17878444	1	92,237,892	DIPplex	InDel	rs17878444	92,237,829	37,518,394	42.3799	0.344906
D1S1627	1	106,963,667	NIST miniplex	STR	rs11260502	106,964,217	14,726,388	11.8331	0.116170
rs3047269	1	162,810,828	HID-38plex	InDel	rs3047269	162,808,601	55,844,384	39.5371	0.329415
D1S1677	1	163,559,721	NIST miniplex	STR	rs11222837	163,559,784	**751.183**	1.8018	0.018011
rs2307924	1	194,878,347	DIPplex	InDel	rs2307924	194,878,370	31,318,586	28.3322	0.256446
D1S1656	1	230,905,307	NGM/Powerplex 16/ES*	core STR	rs113633160	230,905,307	36,026,937	42.7855	0.347020
rs2307579	1	247,812,083	HID-38plex	InDel	rs2307579	247,185,273	16,279,966	33.8562	0.294823
TPOX	2	1,493,413	Identifiler/Powerplex 16/ES*	core STR	rs113475620	1,493,487			
rs2067235	2	41,136,221	DIPplex	InDel	rs2067235	41,133,460	39,639,973	60.7273	0.419019
D2S441	2	68,239,116	NGM/Powerplex 16/ES*	core STR	rs10203882	68,239,020	27,105,560	28.0856	0.254623
rs28369942	2	100,081,561	DIPplex	InDel	rs28369942	100,075,816	31,836,796	24.6366	0.228191
D2S1776	2	169,645,212	NIST miniplex	STR	rs112494037	169,643,383	69,567,567	61.9138	0.422482
rs2307959	2	169,800,273	DIPplex	InDel	rs2307959	169,800,901	**157.518**	0.2630	0.002630
D2S1338	2	218,879,516	Identifiler	STR	rs112111672	218,879,435	49,078,534	46.1908	0.363849
rs16624	2	235,016,391	HID-38plex	InDel	rs16624	235,016,430	16,136,995	23.8073	0.221576
rs2308242	3	8,616,709	HID-38plex	InDel	rs2308242	8,616,181			
D3S1358	3	45,582,207	Identifiler/Powerplex 16/ES*	core STR	rs111694514	45,582,627	36,966,446	45.8002	0.362001
D3S4529	3	85,852,632	NIST miniplex	STR	rs114767567	85,852,702	40,270,075	41.2772	0.339038
D3S3053	3	171,750,874	NIST miniplex	STR	rs111699340	171,749,114	85,896,412	68.6138	0.439605
D4S2408	4	31,304,236	NIST miniplex	STR	rs113424060	31,305,596			
D4S2364	4	93,517,363	NIST miniplex	STR	rs11371557	93,515,918	62,210,322	54.5647	0.398675
rs2308292	4	107,889,773	DIPplex	InDel	rs2308292	107,890,612	14,374,694	10.7479	0.105853
rs2308026	4	119,185,407	HID-38plex	InDel	rs2308026	119,186,180	11,295,568	10.9006	0.107311
FGA	4	155,508,953	Identifiler/Powerplex 16/ES*	core STR	rs67296980	155,508,100	36,321,920	31.0504	0.275915
rs2307526	5	5,125,112	HID-38plex	InDel	rs2307526	5,125,083			
D5S2500	5	58,697,197	NIST miniplex	STR	rs111362704	58,697,354	53,572,271	57.7865	0.409822
rs1160956	5	65,378,460	HID-38plex	InDel	rs1160956	65,377,657	**6,680,303**	3.3150	0.033101
rs1610935	5	66,214,500	DIPplex	InDel	rs1610935	66,214,497	**836.84**	1.0037	0.010035
rs1610937	5	76,745,067	DIPplex	InDel	rs1610937	76,745,024	10,530,527	12.9809	0.126969
D5S818	5	123,111,186	Identifiler/Powerplex 16/ES*	core STR	rs11249749	123,111,652	46,366,628	39.0526	0.326655
CSF1PO	5	149,455,736	Identifiler/Powerplex 16/ES*	core STR	rs113729910	149,455,757	26,344,105	27.7611	0.252212
rs1305056	5	155,662,256	DIPplex	InDel	rs1305056	155,669,910	**6,214,153**	7.0694	0.070227
rs1610871	5	171,087,970	HID-38plex	InDel	rs1610871	171,087,773	15,417,863	21.2861	0.200870
D6S1017	6	41,677,279	NIST miniplex	STR	rs76705065	41,677,034			
rs2307710	6	47,821,263	HID-38plex	InDel	rs2307710	47,821,478	**6,144,444**	9.6850	0.095657
SE33	6	88,986,927	Powerplex 16/ES*	core STR	rs71021371	88,987,046	41,165,568	22.9604	0.214719
rs2307652	6	97,458,121	DIPplex	InDel	rs2307652	97,457,626	**8,470,580**	7.8510	0.077872
D6S474	6	112,879,130	NIST miniplex	STR	rs113991233	112,879,893	15,422,267	15.3622	0.148964
rs2307839	6	117,093,558	HID-38plex	InDel	rs2307839	117,093,929	**4,214,036**	2.3060	0.023043
rs2308137	6	149,614,198	HID-38plex	InDel	rs2308137	149,615,521	32,521,592	34.1515	0.296743
rs2307978	7	83,283,913	HID-38plex	InDel	rs2307978	83,283,614			
D7S820	7	83,789,393	Identifiler/Powerplex 16/ES*	core STR	rs112714641	83,789,257	**505.643**	0.4416	0.004416
rs17879936	7	95,047,150	DIPplex	InDel	rs17879936	95,048,322	11,259,065	8.3214	0.082454
rs1611048	7	110,939,987	DIPplex	InDel	rs1611048	110,938,489	**15,890,167**	15.8230	0.153151
rs1611001	7	154,404,562	DIPplex	InDel	rs1611001	154,403,882	43,465,393	53.0096	0.392869
rs2308072	8	19,089,779	DIPplex	InDel	rs2308072	19,090,033			
D8S1115	8	42,536,587	NIST miniplex	STR	rs113659271	42,546,507	23,456,474	28.6004	0.258416
rs35769550	8	76,518,680	HID-38plex	InDel	rs35769550	76,518,430	33,971,923	25.1881	0.232535
rs589447	8	100,880,861	HID-38plex	InDel	rs589447	100,880,444	24,362,014	17.6300	0.169340
rs3081400	8	119,947,801	DIPplex	InDel	rs3081400	119,947,860	19,067,416	15.9606	0.154398
D8S1179	8	125,907,105	Identifiler/Powerplex 16/ES*	core STR	rs11135824	125,907,272	**5,959,412**	8.2640	0.081896
rs16402	9	38,406,788	HID-38plex	InDel	rs16402	38,407,317			
rs2067294	9	71,314,421	HID-38plex	InDel	rs2067294	71,317,392	32,910,075	7.1213	0.070736

Table 1. (continued)

D9S1122	9	79,688,628	NIST miniplex	STR	rs11283097	79,688,048	**8,370,656**	10.5418	0.103883
rs8190570	9	98,997,911	DIPplex	InDel	rs8190570	99,002,895	19,314,847	23.3166	0.217615
rs2307580	9	105,586,193	HID-38plex	InDel	rs2307580	105,586,672	**6,583,777**	6.2275	0.061955
D9S2157	9	136,035,509	NIST miniplex	STR	rs113499999	135,983,411	30,396,739	45.6616	0.361340
D10S1435	10	2,243,273	NIST miniplex	STR	N/A	2,243,874			
rs140809	10	5,987,163	HID-38plex	InDel	rs140809	5,987,566	**3,743,890**	8.9782	0.088829
rs1160886	10	54,442,386	HID-38plex	InDel	rs1160886	54,443,048	48,455,223	62.3416	0.423697
D10S1248	10	131,092,462	NGM/Powerplex 16/ES*	core STR	rs113518246	131,093,166	76,650,076	95.1776	0.478270
rs10688868	11	268.18	HID-38plex	InDel	rs10688868	2,680,39			
TH01	11	2,192,343	Identifiler/Powerplex 16/ES*	core STR	rs71029110	2,192,549	**1,924,163**	4.3803	0.043691
rs34811743	11	30,177,690	HID-38plex	InDel	rs34811743	30,179,213	27,985,347	42.8315	0.347258
rs17174476	11	102,479,418	DIPplex	InDel	rs17174476	102,478,560	72,301,728	62.1451	0.423141
rs33972805	11	126,288,872	HID-38plex	InDel	rs33972805	126,289,088	23,809,454	32.1409	0.283407
D11S4463	11	130,872,351	NIST miniplex	STR	rs11222421	130,873,262	**4,583,479**	9.5812	0.094656
VWA	12	6,093,104	Identifiler/Powerplex 16/ES*	core STR	rs10579907	6,093,924			
D12S391	12	12,449,950	NGM/Powerplex 16/ES*	core STR	rs113002069	12,449,332	**6,356,846**	11.9408	0.117188
rs1610919	12	14,909,996	HID-38plex	InDel	rs1610919	14,909,486	**2,460,046**	4.1202	0.041109
rs2307570	12	94,675,906	DIPplex	InDel	rs2307570	94,675,496	79,765,910	79.0301	0.459348
D12ATA63	12	108,322,378	NIST miniplex	STR	rs78526997	108,322,352	13,646,472	16.1972	0.156535
rs2067238	12	115,288,548	HID-38plex	InDel	rs2067238	115,288,425	**6,966,170**	8.6381	0.085532
rs17238892	13	31,328,384	DIPplex	InDel	rs17238892	31,328,939			
rs2308171	13	44,880,155	HID-38plex	InDel	rs2308171	44,880,934	13,551,771	18.3500	0.175682
D13S317	13	82,722,079	Identifiler/Powerplex 16/ES*	core STR	rs111980288	82,721,723	37,841,924	32.7842	0.287742
rs2308189	14	29,036,757	HID-38plex	InDel	rs2308189	29,036,632			
rs2308163	14	58,050,081	HID-38plex	InDel	rs2308163	58,050,270	29,013,324	33.1371	0.290092
D14S1434	14	95,308,359	NIST miniplex	STR	rs111914457	95,308,332	37,258,278	43.4457	0.350411
rs2308020	15	53,481,517	HID-38plex	InDel	rs2308020	53,482,122			
rs2307433	15	89,864,316	DIPplex	InDel	rs2307433	89,862,501	36,382,799	45.8764	0.362363
PentaE	15	97,374,392	Powerplex 16/ES*	core STR	rs8036258	97,377,441	**7,510,076**	22.9773	0.214857
rs1610905	16	55,691,830	DIPplex	InDel	rs1610905	55,691,839			
rs2067208	16	84,582,287	HID-38plex	InDel	rs2067208	84,582,117	28,890,457	48.1847	0.372960
D16S539	16	86,386,277	Identifiler/Powerplex 16/ES*	core STR	rs112689398	86,384,543	**1,803,990**	5.8459	0.058194
rs2307581	17	3,970,133	DIPplex	InDel	rs2307581	3,970,395			
rs3051300	17	10,135,941	HID-38plex	InDel	rs3051300	10,135,386	**6,165,808**	15.8217	0.153140
D17S974	17	10,518,733	NIST miniplex	STR	rs112195386	10,518,759	**382.792**	0.6882	0.006881
rs1305047	17	16,084,988	DIPplex	InDel	rs1305047	16,085,364	**5,566,255**	14.3397	0.139590
D17S1301	17	72,680,956	NIST miniplex	STR	rs113771602	72,680,495	56,595,968	71.3992	0.445626
D18S853	18	3,990,543	NIST miniplex	STR	D18S853	3,990,470			
rs3080855	18	23,253,207	HID-38plex	InDel	rs3080855	23,253,013	19,262,664	35.8141	0.307301
rs34511541	18	36,423,040	HID-38plex	InDel	rs34511541	36,423,383	13,169,833	10.6806	0.105211
D18S51	18	60,948,909	Identifiler/Powerplex 16/ES*	core STR	rs10560567	60,949,983	24,525,869	30.3732	0.271169
rs36040336	19	1,402,662	HID-38plex	InDel	rs36040336	1,402,742			
D19S433	19	30,417,028	Powerplex 16/ES*	STR	rs113951851	3,044,163	29,014,366	5.5844	0.055613
rs2307689	19	44,204,340	HID-38plex	InDel	rs2307689	4,422,797	13,787,312	4.5162	0.045039
D20S482	20	4,506,280	NIST miniplex	STR	rs112524392	4,506,638			
rs33917182	20	11,695,625	HID-38plex	InDel	rs33917182	11,696,386	**7,189,345**	17.7336	0.170256
rs16438	20	25,278,470	DIPplex	InDel	rs16438	25,277,915	13,582,845	20.3406	0.192881
rs34541393	20	30,701,405	HID-38plex	InDel	rs7279663	30,701,745	**5,422,935**	0.8779	0.008778
D20S1082	20	53,865,907	NIST miniplex	STR	rs113175620	53,865,700	23,164,502	33.2115	0.290586
rs35605984	21	15,634,865	HID-38plex	InDel	rs35605984	15,635,122			
D21S11	21	20,554,281	Identifiler/Powerplex 16/ES*	core STR	rs113145752	20,554,558	**4,919,416**	10.1116	0.099760
rs10629077	21	31,372,337	HID-38plex	InDel	rs10629077	31,370,907	10,818,056	16.3356	0.157781
rs8177524	21	34,660,756	DIPplex	InDel	rs8178524	34,660,028	**3,288,419**	5.1979	0.051793
PentaD	21	45,056,212	Powerplex 16/ES*	core STR	N/A	45,056,178	10,395,456	23.1968	0.216644
rs16388	22	25,750,816	DIPplex	InDel	rs16388	25,751,136			
rs2307700	22	26,790,901	HID-38plex	InDel	rs2307700	26,790,766	**1,040,085**	3.9666	0.039583
rs6481	22	35,701,900	DIPplex	InDel	rs6481	35,702,663	**8,910,999**	11.5407	0.113400
rs16363	22	37,409,885	DIPplex	InDel	rs16363	37,409,910	**1,707,985**	2.9563	0.029529
D22S1045	22	37,536,318	NGM/Powerplex 16/ES*	core STR	rs112790319	37,535,663	**126.433**	0.0497	0.000497

Several of the analyzed marker pairs display very low Kosambi Rc values, with the lowest being the STR D22S1045 and the DIPplex InDel rs16363, with an Rc of just 0.000497. Therefore these two markers are in very close linkage. It is likely the less stable STR marker will create alleles by slippage or diminution mutations that are likely to remain in very close association with one allele of the more stable binary rs16363, and this association will be broken up by recombination at very slow rates. The situation of other markers on chromosome 22 also represents a high potential for linkage as these five markers lie within 17 Mb of each other. Although the markers DIPplex InDel rs16388 and HID38plex InDel rs2307700 are separated by ~1 Mb they have an Rc value of 0.0395 and therefore can be considered as linked. Higher recombination rates separate two of the markers from the other three with an Rc as high as 0.137. Another DIPplex InDel (rs6481) is only slightly linked to rs16363, a marker included in the D22S1045-rs16363 linkage pair described above, with an Rc of 0.03. It seems that for chromosome 22 only two of the five possible markers could be used as statistically independent loci.

Two of the HID-38plex InDels create pairs with STRs separated by minimal distances, comprising: D7S820-rs2307978 with an Rc of 0.0044 and D12S391-rs1610919 showing 0.0411. Three HID-38plex and DIPplex pairs suggested close linkage, comprising: rs1160956-rs1610935, rs16438-rs34541393 and rs16388-rs2307700 with Rc values of 0.01, 0.0087 and 0.039 respectively. This could be an issue when combining the statistic power of the 68 InDels comprised by the two assays.

Five NIST miniSTRs have been found to be close to certain InDel markers, comprising pairs: D1S1677-rs3047269 (Rc: 0.018), D2S1776-rs2307959 (0.0026), D5S2500-rs1160956 (0.033), D6S474-rs2307839 (0.023), and D17S5974-rs3051300 (0.0068). Of these, the pair D5S2500-rs1160956 on chromosome 5 is also close to the InDel rs1610935, which has a relation to rs1160956 with an Rc value of 0.01 and to D5S2500 with Rc of 0.04201 (value not included in the table) and seemingly less linked to it as the value is only slightly above reference value, although worth of attention.

None of the NIST miniSTRs were found to be in close proximity or exhibiting low Rc values with any other forensic STR with all values above the reference value of 0.04. Therefore each of the STRs assessed in this study can be considered statistically independent and therefore multipliable together.

Taking into account all Rc estimates made between the closest marker pairs, it is apparent that not all InDels in the two forensic assays can be statistically combined with the commonly used STRs. However, only three STR markers within the core are in close linkage with any of the 68 InDels. Even so, InDels in close linkage to core STRs should not be excluded from analyses seeking to combine the allele frequencies into a single probability. As the main application scenario for InDels is typing of degraded DNA, the successful amplification of D7S820, D12SS391 and D22S1045 would not be assured under such conditions when the amplification of InDels could be considered more likely to succeed [1,19,20]. Instead, the information highlighted on this paper should be taken into account to select the marker set with the highest probability to be independent.

Table 2. Observed heterozygosity for Caucasian population for InDel/InDel linkage pairs

Marker	He Cauc
rs1160956	0,238
rs1610935	0,481
rs16438	0,473
rs34541393	0,441
rs16388	0,504
rs2307700	0,487
rs16363	0,519
rs6981	0,416

The same could be concluded if we want to combine the power of the whole 68 InDel markers within those two assays in order to achieve the highest power from both short amplicon PCR assays under highly challenging conditions. If all 68 InDels are successfully amplified, we should be aware that there are four pairs that cannot be treated as statistically independent. Instead it is recommended to select one locus per pair that will be the most informative. In Table 2, we indicate the observed heterozygosity in U.S. Caucasians, calculated from the data obtained in our analysis of United States population samples for HID-38plex and DIPplex [19] (262 independent samples from the U.S. Caucasian

population group). Once we select the marker from each pair with higher heterozygosity, the DIPplex mean Random Match Probability (RMP) for Europeans (removing rs6481), drops from $5.329x10^{-12}$ to $1.0836x10^{-11}$. The HID-38plex mean RMP for Europeans, (removing rs2307978 and rs1610919), drops from $4.7386x10^{-15}$ to $3.1381x10^{-14}$. Combined InDel sets removes four loci (rs6481, rs1160956, rs34541393 and rs2307700), so the RMP drops from $2.53x10^{-26}$ to $8.32x10^{-25}$. Taking into account STRs requires removal of six Indel markers (rs6481, rs1160956, rs34541393, rs2307978, rs1610919 and rs2307700), reducing the RMP to $5.51x10^{-24}$. In all cases even though the assays remain highly informative after compensating for marker linkage, the RMP value is reduced by at least one order of magnitude. Even in the most conservative scenario of taking out all InDel to InDel linked markers and all core STR markers linked to the 68 InDels, the power of the combined marker selection is still comparable to that of 20 combined core STRs [21].

CONCLUSION

We have shown that within the HID-38plex and DIPplex InDel assays there are several markers that are not independent from several core STRs and several other STRs of forensic interest. Some Insertion/Deletion markers are linked to other components within and between the two InDel sets. Far from being a restriction for the incorporation of InDels as commonly applied forensic markers, the information presented in this paper could be applied to make the best use of these highly informative markers in combination with STRs or as a stand-alone set of short amplicon markers. Even after removing all closely linked markers from the calculations, these two InDel assays should provide enough power to help resolving the most challenging DNA cases.

ACKNOWLEDGEMENTS

QIAGEN for providing access to DIPplex kits. Pedro Barrie de la Maza Fundation, for the grant to Manuel Fondevila Alvarez. This work was partially supported through a PhD grant to CS (SFRH/BD/75627/2010) awarded by the Portuguese Foundation for Science and Technology (FCT) and co-financed by the European Social Fund (Human Potential Thematic Operational Programme). The Xunta de Galicia, for the funding associated to the project INCITE 09208163PR granted to M.V. Lareu.

REFERENCES

1- M. Fondevila, C. Phillips, N. Naverán et al*. FSI: Genetics* 2 (2008) 212–218.
2- J. Burger, S. Hummel, B. Herrmann et al*. Electrophoresis* (1999), 20, 1722-1728.
3- C. Romano, E. Di LuiseD. Di Martino et al*. International congress series* 1288 (2006), 577-579.
4- M.D. Coble, J.M. Butler. *J Forensic Sci*, (2005), Vol. 50, No. 1, 43-53.
5- J.J. Sanchez, C. Phillips, C. Borsting et al. *Electrophoresis (*2006), 27, 1713–1724.
6- R. Pereira, C. Phillips, C. Alves et al., *Electrophoresis* (2009) 30: 3682–3690.
7- E. Golenberg, A. Bickel, P. Weihs. *Nucleic Acids Research, (*1996), Vol. 24, No. 24.
8- J.A. Sikorsky, DA. Primerano, TW. Fenger et al. *Biochem. and Biophys. Res. Communications* 355 (2007) 431–437.
9- E.J. Kontanis, FA. Reed. J Forensic Sci, (2006), Vol. 51, No. 4, 795-804.
10- C. Phillips, M. Fondevila, M. Garcia-Magariños. *FSI: Genetics* 2 (2008) 198–204.
11- J.L. Weber, D. David, J. Heil et al*. Am. J. Hum. Genet.* (2002) 71, 854-862*.*
12- J.M Mullaney, R.E. Mills,W.S. Pittard et al*. Hum. Mol. Genet. (2010)* 19(R2): 131-136*.*
13- P. Gill. *Int J Legal Med* (2001) 114: 204–210.
14- WK. Fung, Y. Chung, D. Wong*. Int J Legal Med* (2002) 116 :64–67.
15- C. Phillips, D. Ballard, P. Gill, et al., *FSI: Genetics* (2011) doi:10.1016/fsigen.2011.07.012..
16- http://www.qiagen.com/products/investigatordipplexkit.aspx
17- http://hapmap.ncbi.nlm.nih.gov/downloads/recombination/latest/rates/.
18- C.R. Hill, M.C. Kline, M.D. Coble et al. *J Forensic Sci*, (2008), Vol. 53, No.1. 73-80.
19- C. Romanini, M.L. Catelli, A. Borosky et al. *FSI: Genetics* (2011) doi:10.1016/fsigen.2011.10.006.
20- M. Fondevila, R. Pereira, L. Gusmão et al. *XXIV ISFG meeting poster communication*. (http://www.cstl.nist.gov/strbase/pub_pres/ISFG2011-InDels.pdf).
21- J.M. Butler, C.R. Hill, D.L. Duewer et al. *AAFS 2011 meeting presentation*. (http://www.cstl.nist.gov/strbase/pub_pres/AAFS2011-Butler-additionalSTRloci.pdf).