

7

ISFG DNA Commission on STR Sequence Nomenclature

2016 Consideration	2022 Recommendation
1. Software that allows STR sequences to be exported and stored in databases as sequence strings	1. STRINGS
2. The forward strand direction can be used to align STR sequences	Sequenced STR alleles should be maintained as sequence strings oriented to the forward strand of the current genome assembly
3. GRCh38 is recommended as the framework. Continued discussions are necessary to decide whether or not to adapt to novel genome assemblies	Sequences should include the minimum genomic coordinate range described herein, which is designed to provide sufficient flanking region to distinguish the termini of the repeat region
4. Translate the nomenclature of reverse strand loci and repeat region start and end points.	Recommended minimum range
6. STR sequence strings should include flanking sequences as well as the genome coordinates of the sequence	2. SEQUENCE PROXIES
5. Comprehensive STR nomenclature systems are preferred for early adopters. Backward compatibility	Bracketted repeat and short designator Maintain connection to CE length Universal parameters for all loci
7. Updated allele frequency databases will be necessary	3. RESOURCES
8. Future forensic MPS multiplexes would benefit from retention of past markers	4. NEW LOCI
	5. DATABASING
	<ul style="list-style-type: none"> Length based STR profiles from sequence data can currently be searched/stored Kit information to alert users when sequence data Future database software development

8

Recommendation 1 STRINGS

Sequenced STR alleles should be maintained as sequence strings oriented to the forward strand of the current genome assembly

Sequences should include the minimum genomic coordinate range described herein, which is designed to provide sufficient flanking region to distinguish the termini of the repeat region

9

Recommendation 1 STRINGS

Minimum Reporting Range

How is this useful?

- Exchange of common information
- Algorithmic methods for bracketing and short designators
- Data QC

The full list of sequence strings must be submitted as Supplementary information...sequence strings should include flanking sequences and indicate the genome coordinates of the sequence read.

10

Recommendation 1 STRINGS

Minimum Reporting Range

Minimum Range Rule Set

- 1.1 Allow 5' and 3' shifting of the CE-counted repeat motif to determine the core repeat.
- 1.2 If length variants are observed within +/- 25 bases of the core repeat motif, allow 5' and 3' shifting of the insertion/deletion, repeating 1.1 for these length variations.
- 1.3 SNPs may be present but are not considered in determining the minimum reporting range, except when an SNP is observed which would result in a continuation of the core repeat motif. Include this SNP in the 5' and 3' shifting resulting in an expanded core repeat motif.
- 1.4 Add a four-base cap to each end beyond the core repeat motif.

11

Recommendation 1 STRINGS

Minimum Reporting Range

Minimum Range Rule Set

- 1.1 Allow 5' and 3' shifting of the CE-counted repeat motif to determine the core repeat.

12

Recommendation 4 NEW LOCI

How is this useful?

- ▶ Sequencing allows greater locus multiplexing
- ▶ Researchers are evaluating new STRs
- ▶ Formatting should follow these Recommendations
- ▶ **Historical locus names are difficult to find**
- ▶ **Controls for target fidelity are not available**

Accession	Gene	Chromosome	Position	Size	Repeat	Allele	Database
AF040701	D1S11	18	112,300,000	1.5 kb	10-15	18	STRBase
AF040702	D1S16	18	112,300,000	1.5 kb	10-15	STRBase	
AF040703	D1S19	18	112,300,000	1.5 kb	10-15	STRBase	
AF040704	D1S21	18	112,300,000	1.5 kb	10-15	STRBase	
AF040705	D1S22	18	112,300,000	1.5 kb	10-15	STRBase	
AF040706	D1S23	18	112,300,000	1.5 kb	10-15	STRBase	
AF040707	D1S24	18	112,300,000	1.5 kb	10-15	STRBase	
AF040708	D1S25	18	112,300,000	1.5 kb	10-15	STRBase	
AF040709	D1S26	18	112,300,000	1.5 kb	10-15	STRBase	
AF040710	D1S27	18	112,300,000	1.5 kb	10-15	STRBase	
AF040711	D1S28	18	112,300,000	1.5 kb	10-15	STRBase	
AF040712	D1S29	18	112,300,000	1.5 kb	10-15	STRBase	
AF040713	D1S30	18	112,300,000	1.5 kb	10-15	STRBase	
AF040714	D1S31	18	112,300,000	1.5 kb	10-15	STRBase	
AF040715	D1S32	18	112,300,000	1.5 kb	10-15	STRBase	
AF040716	D1S33	18	112,300,000	1.5 kb	10-15	STRBase	
AF040717	D1S34	18	112,300,000	1.5 kb	10-15	STRBase	
AF040718	D1S35	18	112,300,000	1.5 kb	10-15	STRBase	
AF040719	D1S36	18	112,300,000	1.5 kb	10-15	STRBase	
AF040720	D1S37	18	112,300,000	1.5 kb	10-15	STRBase	
AF040721	D1S38	18	112,300,000	1.5 kb	10-15	STRBase	
AF040722	D1S39	18	112,300,000	1.5 kb	10-15	STRBase	
AF040723	D1S40	18	112,300,000	1.5 kb	10-15	STRBase	
AF040724	D1S41	18	112,300,000	1.5 kb	10-15	STRBase	
AF040725	D1S42	18	112,300,000	1.5 kb	10-15	STRBase	
AF040726	D1S43	18	112,300,000	1.5 kb	10-15	STRBase	
AF040727	D1S44	18	112,300,000	1.5 kb	10-15	STRBase	
AF040728	D1S45	18	112,300,000	1.5 kb	10-15	STRBase	
AF040729	D1S46	18	112,300,000	1.5 kb	10-15	STRBase	
AF040730	D1S47	18	112,300,000	1.5 kb	10-15	STRBase	
AF040731	D1S48	18	112,300,000	1.5 kb	10-15	STRBase	
AF040732	D1S49	18	112,300,000	1.5 kb	10-15	STRBase	
AF040733	D1S50	18	112,300,000	1.5 kb	10-15	STRBase	
AF040734	D1S51	18	112,300,000	1.5 kb	10-15	STRBase	
AF040735	D1S52	18	112,300,000	1.5 kb	10-15	STRBase	
AF040736	D1S53	18	112,300,000	1.5 kb	10-15	STRBase	
AF040737	D1S54	18	112,300,000	1.5 kb	10-15	STRBase	
AF040738	D1S55	18	112,300,000	1.5 kb	10-15	STRBase	
AF040739	D1S56	18	112,300,000	1.5 kb	10-15	STRBase	
AF040740	D1S57	18	112,300,000	1.5 kb	10-15	STRBase	
AF040741	D1S58	18	112,300,000	1.5 kb	10-15	STRBase	
AF040742	D1S59	18	112,300,000	1.5 kb	10-15	STRBase	
AF040743	D1S60	18	112,300,000	1.5 kb	10-15	STRBase	
AF040744	D1S61	18	112,300,000	1.5 kb	10-15	STRBase	
AF040745	D1S62	18	112,300,000	1.5 kb	10-15	STRBase	
AF040746	D1S63	18	112,300,000	1.5 kb	10-15	STRBase	
AF040747	D1S64	18	112,300,000	1.5 kb	10-15	STRBase	
AF040748	D1S65	18	112,300,000	1.5 kb	10-15	STRBase	
AF040749	D1S66	18	112,300,000	1.5 kb	10-15	STRBase	
AF040750	D1S67	18	112,300,000	1.5 kb	10-15	STRBase	
AF040751	D1S68	18	112,300,000	1.5 kb	10-15	STRBase	
AF040752	D1S69	18	112,300,000	1.5 kb	10-15	STRBase	
AF040753	D1S70	18	112,300,000	1.5 kb	10-15	STRBase	
AF040754	D1S71	18	112,300,000	1.5 kb	10-15	STRBase	
AF040755	D1S72	18	112,300,000	1.5 kb	10-15	STRBase	
AF040756	D1S73	18	112,300,000	1.5 kb	10-15	STRBase	
AF040757	D1S74	18	112,300,000	1.5 kb	10-15	STRBase	
AF040758	D1S75	18	112,300,000	1.5 kb	10-15	STRBase	
AF040759	D1S76	18	112,300,000	1.5 kb	10-15	STRBase	
AF040760	D1S77	18	112,300,000	1.5 kb	10-15	STRBase	
AF040761	D1S78	18	112,300,000	1.5 kb	10-15	STRBase	
AF040762	D1S79	18	112,300,000	1.5 kb	10-15	STRBase	
AF040763	D1S80	18	112,300,000	1.5 kb	10-15	STRBase	
AF040764	D1S81	18	112,300,000	1.5 kb	10-15	STRBase	
AF040765	D1S82	18	112,300,000	1.5 kb	10-15	STRBase	
AF040766	D1S83	18	112,300,000	1.5 kb	10-15	STRBase	
AF040767	D1S84	18	112,300,000	1.5 kb	10-15	STRBase	
AF040768	D1S85	18	112,300,000	1.5 kb	10-15	STRBase	
AF040769	D1S86	18	112,300,000	1.5 kb	10-15	STRBase	
AF040770	D1S87	18	112,300,000	1.5 kb	10-15	STRBase	
AF040771	D1S88	18	112,300,000	1.5 kb	10-15	STRBase	
AF040772	D1S89	18	112,300,000	1.5 kb	10-15	STRBase	
AF040773	D1S90	18	112,300,000	1.5 kb	10-15	STRBase	
AF040774	D1S91	18	112,300,000	1.5 kb	10-15	STRBase	
AF040775	D1S92	18	112,300,000	1.5 kb	10-15	STRBase	
AF040776	D1S93	18	112,300,000	1.5 kb	10-15	STRBase	
AF040777	D1S94	18	112,300,000	1.5 kb	10-15	STRBase	
AF040778	D1S95	18	112,300,000	1.5 kb	10-15	STRBase	
AF040779	D1S96	18	112,300,000	1.5 kb	10-15	STRBase	
AF040780	D1S97	18	112,300,000	1.5 kb	10-15	STRBase	
AF040781	D1S98	18	112,300,000	1.5 kb	10-15	STRBase	
AF040782	D1S99	18	112,300,000	1.5 kb	10-15	STRBase	
AF040783	D1S100	18	112,300,000	1.5 kb	10-15	STRBase	


Genome in a Bottle (GIAB)

19

Recommendation 5 DATABASING

How is this useful?

- ▶ Provides *food for thought*
- ▶ Emphasizes current ability to convert STR sequences into length-based STR profiles
- ▶ Discusses adaptations of current databases to STR sequences
 - ▶ e.g. including kit information
- ▶ Describes future database design considerations



20

Thank You!


STRAND Working Group
 Jonathan King – UNTHSC
 Lisa Borzuck – NIST
 Chris Phillips – USC
 Martin Bodner – MUI
 David Ballard – KCL
 Walther Parson – MUI

Additional Commission Members & Contributors
 Claus Barsting
 Bruce Budowle
 Peter de Krijff
 Peter Gill
 Douglas Hares
 Jerry Hoogenboom
 Jodi Irwin
 Kris van der Gaag
 Sascha Willuweit

NIST Applied Genetics Group

ISFG Executive Board
 John Butler
 Leonor Gusmao
 Peter Schneider
 Marieke Vennemann

Funding Sources
 NIST Special Programs Office
 NJ Interagency Agreement



21