HML Data Dictionary

Table of Contents

Table of Contents

HML Data Dictionary

The purpose of this document is to facilitate the clear understanding of Histoimmunogenetics Markup Language (HML), an Extensible Markup Language (XML) format designed for exchange of human leukocyte antigen (HLA) testing results and primary data. The document attempts to provide a single convenient and comprehensive reference, with the goal of reducing the need to locate and analyze other information sources.

The first few sections offer an overview and comparison of several versions of HML, including version 0.3.3, version 0.3, and version 0.2. These sections use block diagrams to illustrate basic HML document structures.

A *Specialized Data Types* section describes specific data types used within HML, such as DNA sequence and fully qualified test name.

An *Alphabetical Tag Reference* section describes each HML XML element in detail.

A *References* section lists external references.

A series of appendixes show the XML Document Type Definition (DTD) for each HML version.

A brief HML example (HML version 0.3).

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE hml PUBLIC "-//NMDP//DTD HML 0.3//EN" "http://www.nmdp.org/DTD/hml-0.3.dtd">
<hml version="0.3">
  <project name="ABDR">
    <reporting-center code="999"/>
    <sample id="444-333-2" center-code="123">
      <typing>
        <interpretation date="20040123">
          <haploid locus="A" method="DNA" type="01BC"/>
          <haploid locus="A" method="DNA" type="01CF"/>
        </interpretation>
      </typing>
    </sample>
  </project>
</hml>
```

HML Version 0.3.3

The block diagram below shows a high-level overview of the HML version 0.3.3 document structure. Grey shading indicates an optional element, and ellipses (**...**) indicate a repeatable element.

The HML document consists of an **hml** element containing zero or more **property** elements and one or more **project** elements. Each **project** element contains a **reporting-center** element, one or more **sample** elements, and zero or more **typing-test-names** elements.

hml

> *property ...*

> *project*
>
>> *reporting-center*
>>
>> *sample*
>>
>> *sample ...*
>>
>> *typing-test-names ...*

> *project ...*
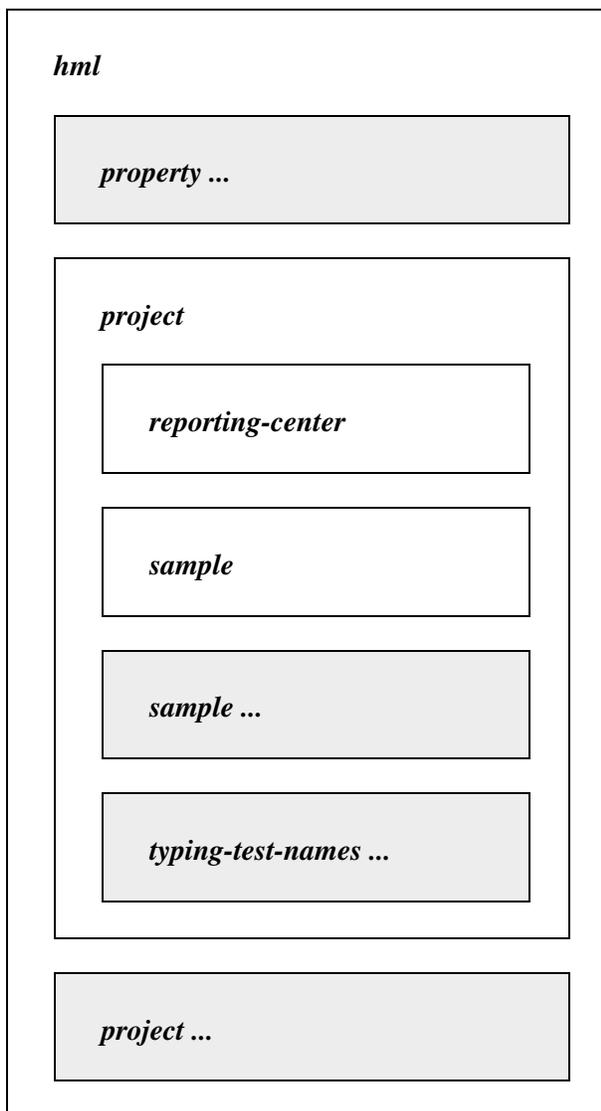
A *sample* element contains zero or more *property* elements and one or more *typing* elements.

```
┌─────────────────────────────────────────┐
│  sample                                  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   property ...                    │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   typing                          │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   typing ...                      │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
└─────────────────────────────────────────┘
```
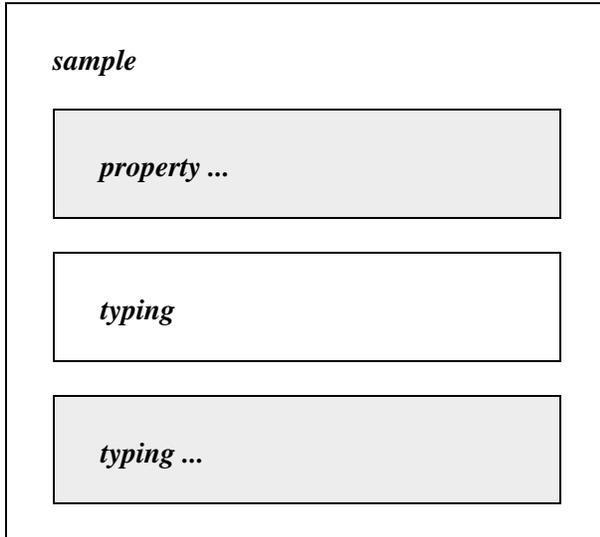
A *typing* element may contain an *interpretation* element and zero or more *sso*, *ssp*, or *sbt* elements.

```
┌─────────────────────────────────────────┐
│  typing                                  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   interpretation                  │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   sso or ssp or sbt ...           │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
└─────────────────────────────────────────┘
```
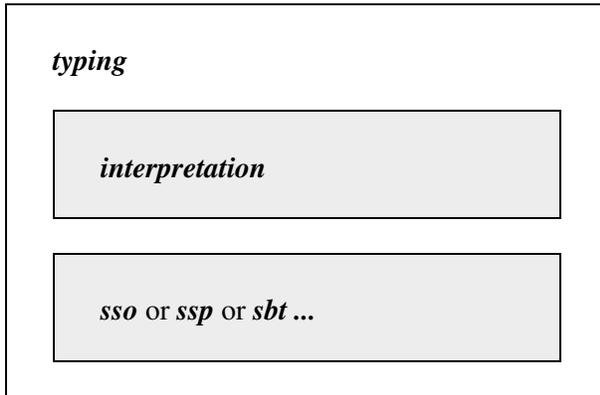
The *interpretation* element has two possible formats. It may contain either *haploid* or *genotype-list* child elements, but not both.

In one of its possible formats, an *interpretation* element contains one or more *haploid* elements.

> **interpretation**
>
> > **haploid**
> >
> > ---
> >
> > **haploid ...**
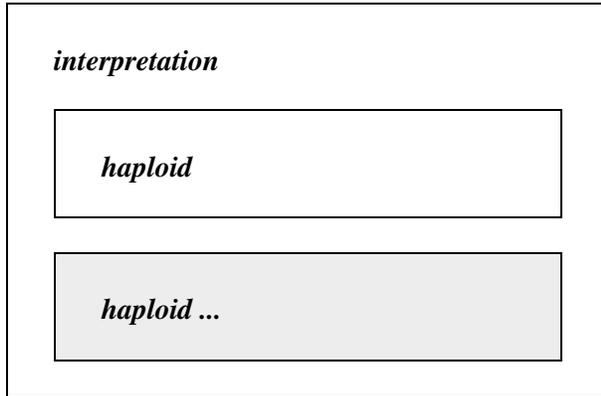
In its other possible format, an *interpretation* element contains one or more *genotype-list* elements. Each *genotype-list* element contains one or more *diploid-combination* elements.

> **interpretation**
>
> > **genotype-list**
> >
> > > **diploid-combination**
> > >
> > > ---
> > >
> > > **diploid-combination ...**
> >
> > **genotype-list ...**
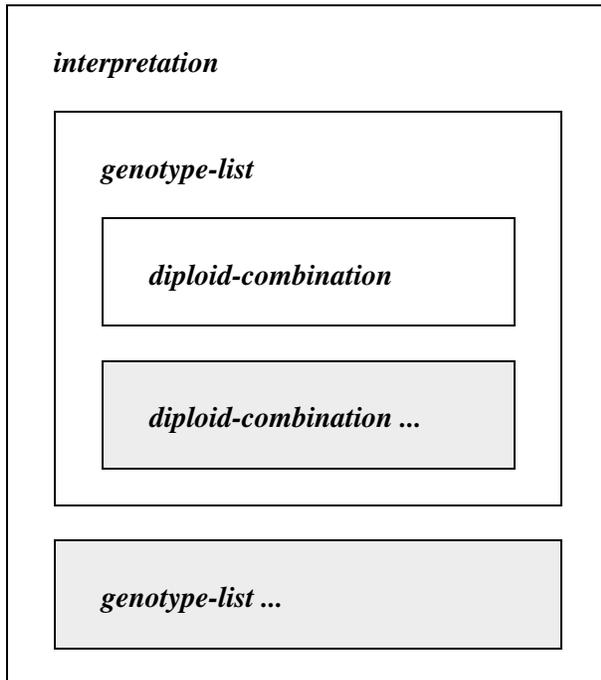
A *diploid-combination* element contains one or two *locus-block* elements. Each *locus-block* element contains one or more *allele-list* elements, and each *allele-list* contains one or more *allele* elements.
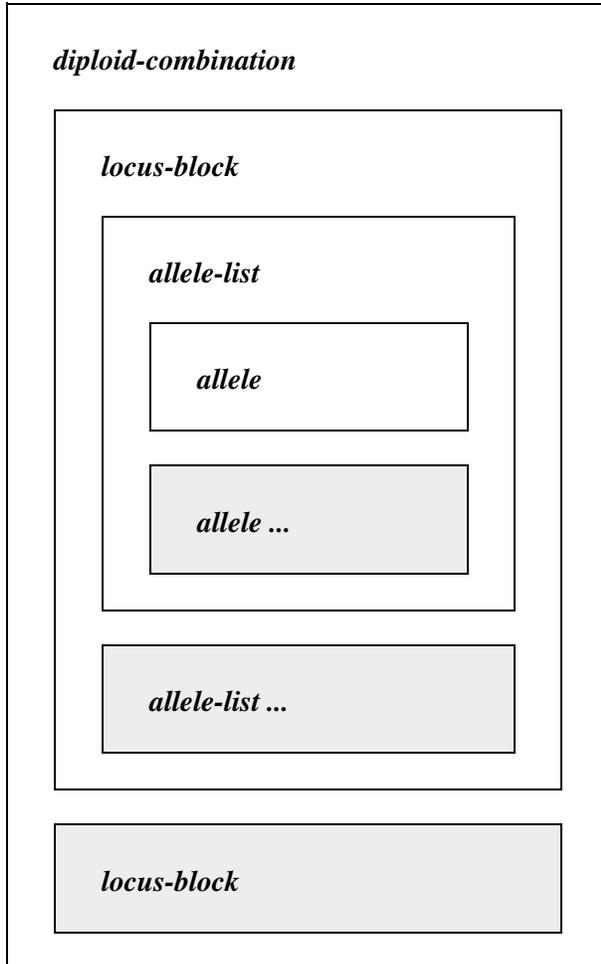
A *sbt* element may contain one *amplification-primer* and one *gssp* element, and must contain one *sequence* element.

```
┌─────────────────────────────────────────────┐
│  sbt                                          │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   amplification-primer                │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   gssp                                │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   sequence                            │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
└─────────────────────────────────────────────┘
```

An *amplification-primer* element contains one *registered-name* element and one *score* element.

```
┌─────────────────────────────────────────────┐
│  amplification-primer                         │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   registered-name                     │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   score                               │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
└─────────────────────────────────────────────┘
```

The *gssp* element has two possible formats. It may either contain a *sequence* child element or a *registered-name* child element, but not both.

In one of its possible formats, a *gssp* element contains one *sequence* element.

```
┌─────────────────────────────────────────────┐
│  gssp                                         │
│    ┌───────────────────────────────────────┐ │
│    │                                       │ │
│    │   sequence                            │ │
│    │                                       │ │
│    └───────────────────────────────────────┘ │
└─────────────────────────────────────────────┘
```
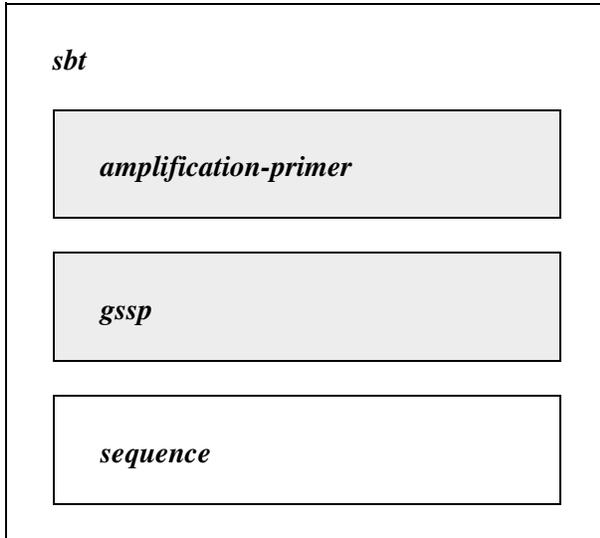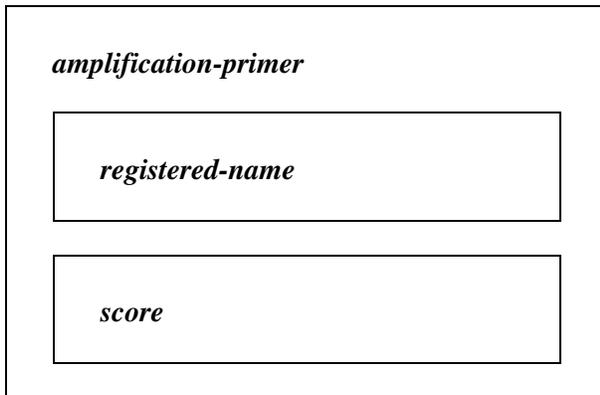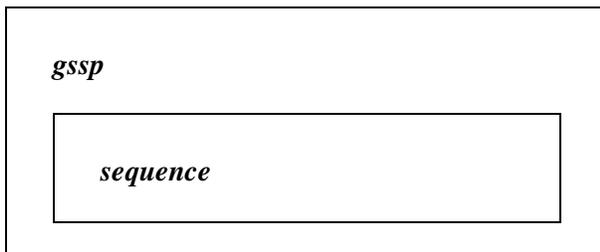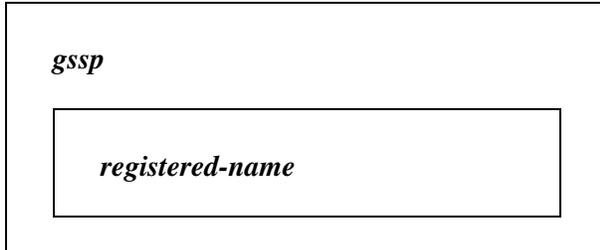
In its other possible format, a *gssp* element contains one *registered-name* element.

```
gssp

    registered-name
```

A *typing-test-names* element contains one or more *typing-test-name* elements.

```
typing-test-names

    typing-test-name

    typing-test-name ...
```
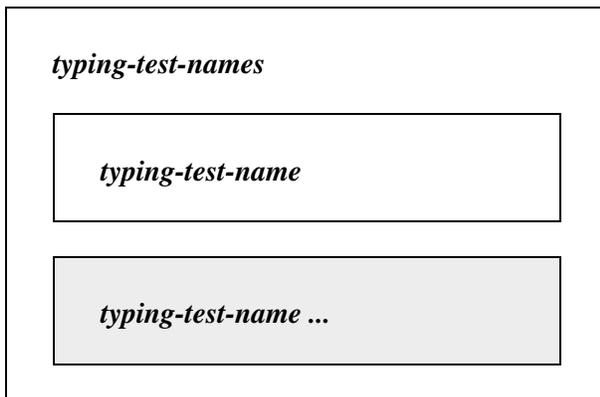
HML Version 0.3

Except for the *sbt* element, HML version 0.3 is the same as version 0.3.3.

The *HML 0.3 sbt* element contains one **sequence** element. It uses *name* and *score* attributes for the test name and primer amplification results, but has no means for describing a Group Specific Sequencing Primer (GSSP).

```
┌─────────────────────────────────────────┐
│  sbt                                      │
│                                           │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │    sequence                       │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
│                                           │
└─────────────────────────────────────────┘
```
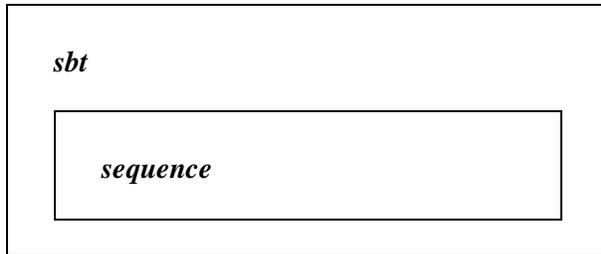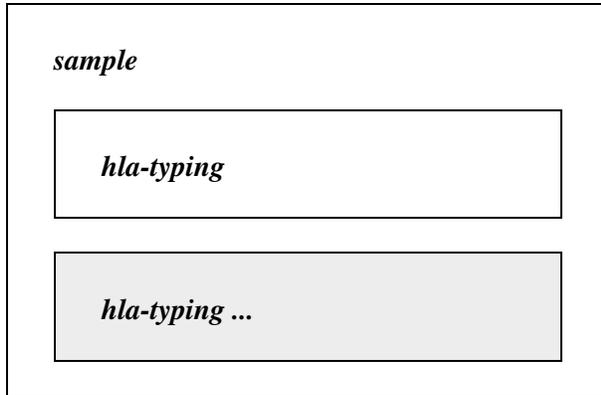
HML Version 0.2

There are several notable differences between HML version 0.2 and HML version 0.3.

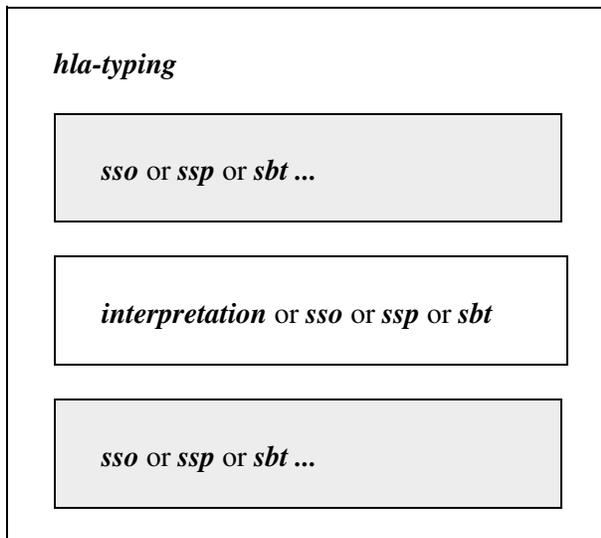The *HML 0.2 **hml*** element contains no ***property*** child elements.

```
hml
    project
        reporting-center

        sample

        sample ...

        typing-test-names ...

    project ...
```
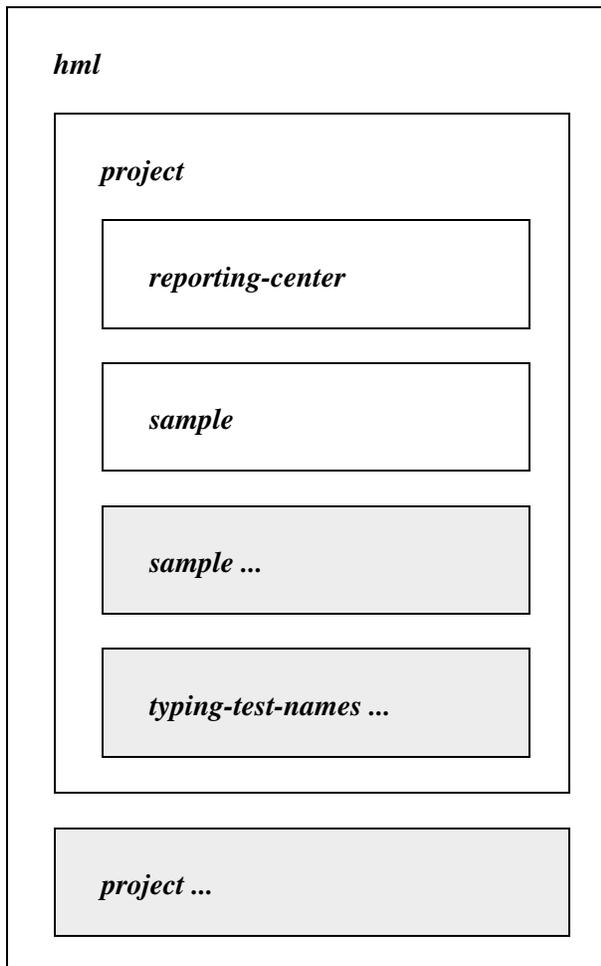
The *HML 0.2 sample* element contains one or more **hla-typing** child elements (instead of **typing** child elements), and no **property** child elements.
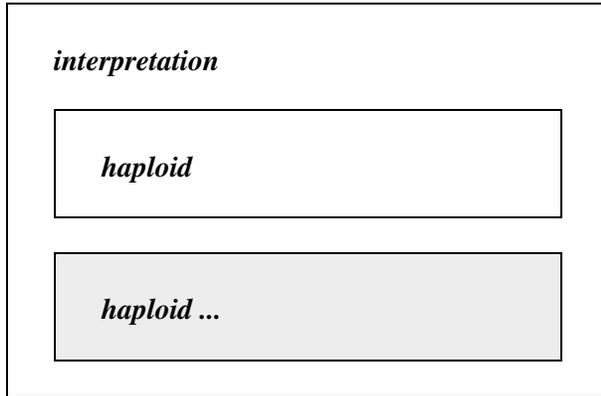
```
┌─────────────────────────────────────────────┐
│  sample                                      │
│                                              │
│   ┌──────────────────────────────────────┐   │
│   │                                      │   │
│   │    hla-typing                        │   │
│   │                                      │   │
│   └──────────────────────────────────────┘   │
│                                              │
│   ┌──────────────────────────────────────┐   │
│   │                                      │   │
│   │    hla-typing ...                    │   │
│   │                                      │   │
│   └──────────────────────────────────────┘   │
│                                              │
└─────────────────────────────────────────────┘
```

A **hla-typing** element contains one **interpretation**, **sso**, **ssp**, or **sbt** element, and zero or more additional **sso**, **ssp**, or **sbt** elements.

```
┌─────────────────────────────────────────────┐
│  hla-typing                                  │
│                                              │
│   ┌──────────────────────────────────────┐   │
│   │    sso or ssp or sbt ...             │   │
│   └──────────────────────────────────────┘   │
│                                              │
│   ┌──────────────────────────────────────┐   │
│   │                                      │   │
│   │    interpretation or sso or ssp or sbt │ │
│   │                                      │   │
│   └──────────────────────────────────────┘   │
│                                              │
│   ┌──────────────────────────────────────┐   │
│   │                                      │   │
│   │    sso or ssp or sbt ...             │   │
│   │                                      │   │
│   └──────────────────────────────────────┘   │
│                                              │
└─────────────────────────────────────────────┘
```
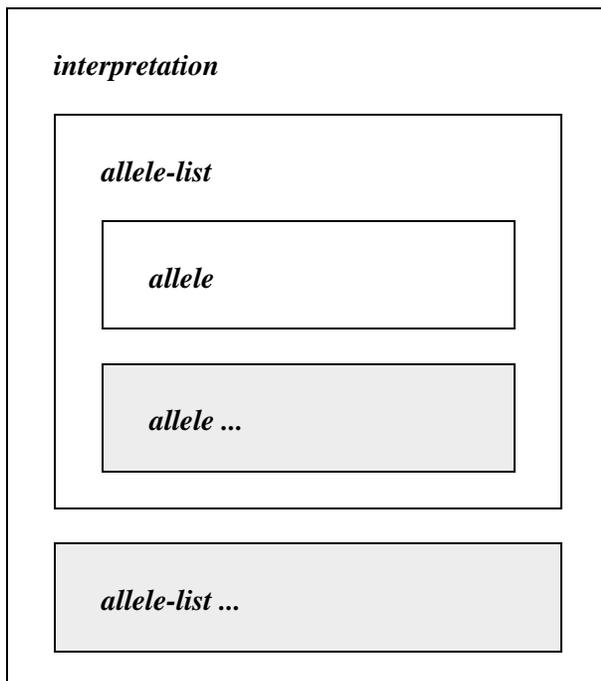
The *HML 0.2 interpretation* element has two possible formats. It may contain **haploid** or **allele-list** child elements, but not both.

In one of its possible formats, a HML 0.2 *interpretation* element contains one or more **haploid** elements.

```
┌─────────────────────────────────────────┐
│ interpretation                           │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   haploid                         │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   haploid ...                     │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
└─────────────────────────────────────────┘
```

In its other possible format, a *HML 0.2 interpretation* element contains one or more **allele-list** elements. Each **allele-list** element contains one or more **allele** elements.

```
┌─────────────────────────────────────────┐
│ interpretation                           │
│   ┌───────────────────────────────────┐  │
│   │ allele-list                       │  │
│   │   ┌─────────────────────────────┐ │  │
│   │   │                             │ │  │
│   │   │   allele                    │ │  │
│   │   │                             │ │  │
│   │   └─────────────────────────────┘ │  │
│   │   ┌─────────────────────────────┐ │  │
│   │   │                             │ │  │
│   │   │   allele ...                │ │  │
│   │   │                             │ │  │
│   │   └─────────────────────────────┘ │  │
│   └───────────────────────────────────┘  │
│   ┌───────────────────────────────────┐  │
│   │                                   │  │
│   │   allele-list ...                 │  │
│   │                                   │  │
│   └───────────────────────────────────┘  │
└─────────────────────────────────────────┘
```

Specialized Data Types

NMDP ID

The NMDP ID is a numeric value which identifies a specific National Marrow Donor Program (NMDP) donor or recipient. In text format, these are typically zero-padded and hyphenated (e.g. "0123-4567-8", "001-234-5"). In HML, zero-padding is not required, and hyphens should be omitted (e.g. "1234567", "12345").

Center Code

The center code is a NMDP donor center or transplant center number.

Fully Qualified Test Name

The fully qualified test name is a character string which identifies a specific SSO or SSP probe kit, SBT test, or GSSP sequence (e.g. "L747.K1.V1.SBTHLAA"). The name format uses dots and letters to separate a series of values describing the test.

"L" + *cl_cde* + **".K"** + *kit_seq* + **".V"** + *version* + **"."** + *name*

For example, shown below is a database query which constructs a fully qualified test name (a.k.a. `kit_name`).

```
select kit_id, method, class, locus
    , 'L' + convert(varchar(5), cl_cde)
    + '.K' + convert(varchar(3), kit_seq)
    + '.V' + convert(varchar(3), version)
    + '.' + name kit_name
from mdp_cmn_prd..t_typing_kit
order by method, class, locus
```

Locus Name

Locus name is a character string that identifies a specific HLA locus, such as "A", "B", "C", "DPA1", "DQA1", "DQB1", or "DRB1".

Typing Method

Typing method is a character string that identifies the method used, such as "DNA" or "SEROLOGY".

Allele Nomenclature

The European Bioinformatics Institute (EBI) maintains the IMGT/HLA Database, which defines the allele nomenclature used by HML. The EBI web site contains additional information about this database.

http://www.ebi.ac.uk/imgt/hla/

Within an *interpretation*, the *genotype-list* element should specify the HLADB version which was used for that interpretation. Example:

```
<interpretation date="20110101">
  <genotype-list db="HLADB" version="3.4.0">
    <diploid-combination>
      <locus-block>
        <allele-list>
          <allele>DRB1*01:01</allele>
        </allele-list>
      </locus-block>
      <locus-block>
        <allele-list>
          <allele>DRB1*01:02</allele>
        </allele-list>
      </locus-block>
    </diploid-combination>
  </genotype-list>
</interpretation>
```

Within an *interpretation* containing *haploid* elements, it may be necessary to use NMDP Allele Codes to describe the interpretation with sufficient accuracy. Example:

```
<interpretation date="20110101">
  <haploid locus="A" method="DNA" type="03:AGWV"/>
  <haploid locus="A" method="DNA" type="26:AHTR"/>
</interpretation>
```

The NMDP Bioinformatics web site contains additional information about NMDP Allele Codes.

http://bioinformatics.nmdp.org/

Test Score

The score associated with a test should be either "1" or "8", where "8" indicates amplification occurred and "1" indicates no amplification occurred. For a probe kit, the scores string contains a series of such scores, e.g. "8181111181818111111".

If the primary data (e.g. *sso* or *ssp*) uses *ref-id* to reference *typing-test-names* construct, there is a one-to-one mapping between scores and typing tests, so the number of scores given must match the corresponding number of *typing-test-name* child elements.

For example, this ***sso*** element uses *ref-id* to reference a ***typing-test-names*** element appearing later within the same HML document.

```
<sso ref-id="ABC123" scores="8818"/>
```

In this example, the above ***sso*** *scores* map to tests specified by ***typing-test-name*** child elements within the referenced ***typing-test-names*** element.

```
<typing-test-names ref-id="ABC123">
  <typing-test-name name="DRB61WNSQK"/>
  <typing-test-name name="DRB9WQLKFE"/>
  <typing-test-name name="DRB10-P-RE"/>
  <typing-test-name name="DRB9EYSTSE"/>
</typing-test-names>
```

DNA Sequence

DNA sequences are character strings composed of nucleic acid codes (e.g. "GCTCCCACTCCATG"), as described below.

IUPAC Nucleic Acid Codes

HML uses letter codes designated by the *International Union of Pure and Applied Chemistry* (IUPAC) to represent ambiguity in nucleic acid sequences [3].

Nucleotide Code	Base
A	adenine
C	cytosine
G	guanine
T	thymine
R	A or G (purine)
Y	C or T (pyrimidine)
K	G or T (keto)
M	A or C (amino)
S	C or G (strong)
W	A or T (weak)
B	C, G, or T (not A)
D	A, G, or T (not C)
H	A, C, or T (not G)
V	A, C, or G (not T)
N	A, C, G, or T (any)

Notes

1. A HML sequence may contain whitespace characters such as space, tab, carriage return, and line feed. However, the presence of whitespace characters has no effect upon the meaning of the sequence, which must be interpreted as if no whitespace characters were present.
2. A HML sequence should not use the code 'X' to represent an unknown nucleotide. For an unknown nucleotide, the code 'N' is preferred.
3. A HML sequence should not use lower case letters.
4. A HML sequence must not use the codes '-', or '.' to represent an unknown nucleotide or sequence gap.
5. A HML sequence must not use other accessory symbols, such as '?', '+', '/', '=', numerals, or lower-case letters.
6. HML sequences should not use the code 'U' (uracil) because they describe DNA, not RNA.

Alphabetical Tag Reference

<allele>

The *allele* element specifies a single allele.

Attribute	Usage	Status	Description
source	*HML 0.2*	required	Namespace identifier for the name attribute. This can be thought of as the HLA sequence database release version (e.g. "HLADB-2.0.0").
name	*HML 0.2*	required	Identifier for the allele (the allele name, e.g. "A*010101").
present	*HML 0.3+*	optional	Indicates presence or absence of the allele. "N" indicates testing to detect the given allele was performed and the allele was found not to be a possibility. "U" (untested) indicates testing to detect the given allele was not performed. The default value is "Y", which indicates testing detected the allele.

In *HML 0.2*, the **allele** element contains only *source* and *name* attributes.

As of *HML 0.3*, the *source* and *name* attributes were eliminated, and the **allele** element itself became a container for a character value which specifies the allele name in `LOCUS*NAME` format. The allele name must always be at allele-level resolution, identifying a single allele. Example:

```
<allele>DRB1*1206</allele>
```

If the **allele** element appears within a structure enclosed by a **genotype-list** element which specifies a nomenclature, the **allele** value must conform to that nomenclature. Otherwise, the **allele** value should follow a nomenclature fully understood by the document recipient.

Change History for <allele>

HML Version	Change Description
HML 0.3	Removed *source* attribute. Removed *name* attribute. Added *present* attribute. Changed **allele** element to container for character value (allele name).

<allele-list>

The *allele-list* element is a representation of the list of allele possibilities for a type.

Child Element	Usage	Status	Multiplicity	Notes
allele	*HML 0.2+*	required	1 or more	

The NMDP has historically used allele codes in combination with allele families. For example, `A*01BC` can be represented as follows (this example uses *HML 0.3* style *allele* elements).

```
<allele-list>
    <allele>A*01:02</allele>
    <allele>A*01:03</allele>
</allele-list>
```

Change History for <allele-list>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<amplification-primer>

The *amplification-primer* element identifies an amplification primer and indicates the result of its use.

Child Element	Usage	Status	Multiplicity	Notes
registered-name	*HML 0.3.3+*	required	1	
score	*HML 0.3.3+*	required	1	

Change History for <amplification-primer>

HML Version	Change Description
HML 0.3.3	Introduced *amplification-primer* element.

<diploid-combination>

The *diploid-combination* element represents one possibility value within a *genotype-list*.

Child Element	Usage	Status	Multiplicity	Notes
locus-block	*HML 0.3+*	required	1 or 2	

The *diploid-combination* element contains either one or two *locus-block* child elements, depending on whether it holds data for one or two chromosomes.

Change History for <diploid-combination>

HML Version	Change Description
HML 0.3	Introduced *diploid-combination* element.

<genotype-list>

The *genotype-list* element represents the full unambiguous list of possibilities for the typing of a sample.

Attribute	Usage	Status	Description
db	HML 0.3+	optional	Name of database or other source which defines the nomenclature for *allele* values appearing within the *genotype-list* (e.g. "HLADB").
version	HML 0.3+	optional	Specific version of the source identified by the *db* attribute (e.g. "2.8.0").

Child Element	Usage	Status	Multiplicity	Notes
diploid-combination	HML 0.3+	required	1 or more	

If *db* and *version* attributes are specified, *allele* values contained within the *genotype-list* must conform to the specified nomenclature. Otherwise, enclosed *allele* values should follow a nomenclature fully understood by the document recipient.

Change History for <genotype-list>

HML Version	Change Description
HML 0.3	Introduced *genotype-list* element.

<gssp>

The *gssp* element describes a Group Specific Sequencing Primer (GSSP).

Child Element	Usage	Status	Multiplicity	Notes
sequence	HML 0.3.3+	optional	1	ambiguous nucleotide codes not allowed
registered-name	HML 0.3.3+	optional	1	

One and only one of *sequence* or *registered-name* must be specified. Use *sequence* to embed a GSSP sequence within the document, or use *registered-name* to reference a sequence already registered with the document recipient.

Change History for <gssp>

HML Version	Change Description
HML 0.3.3	Introduced *gssp* element.

<haploid>

The *haploid* element specifies one-half of a full typing at a particular locus.

Attribute	Usage	Status	Description
locus	*HML 0.2+*	required	Locus (e.g. "A", "DRB1")
method	*HML 0.2+*	required	Typing method used (e.g. "DNA", "SEROLOGY")
type	*HML 0.2+*	required	Allele/code level type (e.g. "0101", "01AB")

Change History for <haploid>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<hla-typing>

The *hla-typing* element encapsulates a typing call, along with primary data (*sso*, *ssp*, and/or *sbt* child elements) which may have been used to determine that call.

Child Element	Usage	Status	Multiplicity	Notes
interpretation	*HML 0.2*	optional	0 or 1	
sso	*HML 0.2*	optional	0 or more	
ssp	*HML 0.2*	optional	0 or more	
sbt	*HML 0.2*	optional	0 or more	

At least one *interpretation*, *sso*, *ssp*, or *sbt* child element must be specified.

Change History for <hla-typing>

HML Version	Change Description
HML 0.3	Eliminated ***hla-typing*** element.

<hml>

The ***hml*** element is the root element of the document.

Attribute	Usage	Status	Description
version	*HML 0.2+*	required	HML version used by the document

Child Element	Usage	Status	Multiplicity	Notes
property	*HML 0.3+*	optional	0 or more	
project	*HML 0.2+*	required	1 or more	

As of *HML 0.3*, document level properties may be specified here using ***property*** child elements.

Change History for <hml>

HML Version	Change Description
HML 0.3	Added ***property*** child element.

<interpretation>

The ***interpretation*** element specifies the typing call at allele/code level.

Attribute	Usage	Status	Description
date	*HML 0.2+*	required	Date on which the typing was computed, or on which the final call was determined. Format can be either ISO 8601 or "YYYYMMDD" format (*HML 0.2* requires "YYYYMMDD" format).

Child Element	Usage	Status	Multiplicity	Notes
haploid	*HML 0.2+*	optional	0 or more	
allele-list	*HML 0.2*	optional	0 or more	
genotype-list	*HML 0.3+*	optional	0 or more	

In *HML 0.2*, the ***interpretation*** element must contain at least one ***haploid*** or ***allele-list*** child element. If it contains any ***haploid*** child elements, it must not contain any ***allele-list*** child elements, and vice versa.

Similarly, as of *HML 0.3*, the ***interpretation*** element must contain at least one ***haploid*** or ***genotype-list*** child element. If it contains any ***haploid*** child elements, it must not contain any ***genotype-list*** child elements, and vice versa.

Change History for <interpretation>

HML Version	Change Description
HML 0.3	Revised *date* attribute to support ISO 8601 date format. Removed ***allele-list*** child element. Added ***genotype-list*** child element.

<locus-block>

The ***locus-block*** element groups ***allele-list*** elements together, indicating one ***allele-list*** is a possibility if and only if all others are.

Child Element	Usage	Status	Multiplicity	Notes
allele-list	*HML 0.3+*	optional	1 or more	

The ability to group related ***allele-list*** elements within a ***locus-block*** is useful, for example, when listing DRB1 alleles adjacent to corresponding DRB3 alleles that are relevant only in some cases. Example:

```
<locus-block>
    <allele-list>
        <allele>DRB1*03:01:01</allele>
        <allele>DRB1*03:03</allele>
    </allele-list>
    <allele-list>
        <allele>DRB3*02:03</allele>
    </allele-list>
</locus-block>
```

Change History for <locus-block>

HML Version	Change Description
HML 0.3	Introduced ***locus-block*** element.

<project>

The *project* element groups data into sections with a common destination or purpose.

Attribute	Usage	Status	Description
name	*HML 0.2+*	required	identifier for data's intent (e.g. "STAR", "SG09")

Child Element	Usage	Status	Multiplicity	Notes
reporting-center	*HML 0.2+*	required	1	
sample	*HML 0.2+*	required	1 or more	
typing-test-names	*HML 0.2+*	optional	0 or more	

The *name* attribute denotes the destination or purpose of the *project*.

Change History for <project>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<property>

The *property* element is a generic name-value pair.

Attribute	Usage	Status	Description
name	*HML 0.3+*	required	
value	*HML 0.3+*	required	

Change History for <property>

HML Version	Change Description
HML 0.3	Introduced *property* element.

<registered-name>

The *registered-name* element contains an identifying name that should be understood by the document recipient. The element has no attributes, and acts only as a container for its value. Example:

```
<registered-name>L999.K1.V1.SBT54321</registered-name>
```

In HML sent to the NMDP, the *registered-name* value is typically the fully qualified test name.

Change History for <registered-name>

HML Version	Change Description
HML 0.3.3	Introduced *registered-name* element.

<reporting-center>

The *reporting-center* element specifies the origin of the data.

Attribute	Usage	Status	Description
code	*HML 0.2+*	required	reporting center code identifier

Change History for <reporting-center>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<sample>

The *sample* element encloses the typing data pertaining to a particular sample.

Attribute	Usage	Status	Description
id	*HML 0.2+*	required	NMDP ID in non-hyphenated format. Need not be zero-padded.
center-code	*HML 0.2+*	required	Center code of the sample's origin (e.g. donor center, transplant center).

Child Element	Usage	Status	Multiplicity	Notes
hla-typing	*HML 0.2*	required	1 or more	
property	*HML 0.3+*	optional	0 or more	
typing	*HML 0.3+*	required	1 or more	

As of *HML 0.3*, sample properties may be specified here using *property* child elements.

Change History for <sample>

HML Version	Change Description
HML 0.3	Removed **hla-typing** child element. Added **property** child element. Added **typing** child element.

<sbt>

The **sbt** element describes a sequence-based typing (SBT) test that was performed. Example (HML 0.3.3, with **amplification-primer** and **gssp**):

```
<sbt>
  <amplification-primer>
    <registered-name>L999.K1.V1.SBT54321</registered-name>
    <score>8</score>
  </amplification-primer>
  <gssp>
    <sequence>ACGTACGT</sequence>
  </gssp>
  <sequence>ACGTACGTACGTRYKMSW</sequence>
</sbt>
```

Attribute	Usage	Status	Description
name	*HML 0.2* *HML 0.3*	required	Fully qualified name of the SBT test (e.g. "L999.K1.V1.SBT54321")
score	*HML 0.2* *HML 0.3*	required	Result of primer amplification. This is usually '8', but could be '1', indicating no amplification happened and no sequence is likely to have been obtained.

Child Element	Usage	Status	Multiplicity	Notes
amplification-primer	*HML 0.3.3+*	optional	0 or 1	
gssp	*HML 0.3.3+*	optional	0 or 1	
sequence	*HML 0.2+*	required	1	ambiguous nucleotide codes allowed

In *HML 0.2* and *HML 0.3*, the *name* and *score* attributes specify the test name and primer amplification result.

As of *HML 0.3.3*, an optional **amplification-primer** child element may contain the test name and primer amplification result. Additionally, *HML 0.3.3* adds a **gssp** child element which can describe a Group Specific Sequencing Primer (GSSP).

Change History for <sbt>

HML Version	Change Description
HML 0.3.3	Removed *name* attribute. Removed *score* attribute. Added **amplification-primer** child element. Added **gssp** child element.

<score>

The *score* element contains the result of primer amplification. The element has no attributes, and acts only as a container for its value. Example:

```
<score>8</score>
```

The score value is usually '8', but could be '1', indicating no amplification happened and no sequence is likely to have been obtained.

Change History for <score>

HML Version	Change Description
HML 0.3.3	Introduced *score* element.

<sequence>

The *sequence* element describes a DNA sequence. It has no attributes, and acts only as a container for its value. Example:

```
<sequence>CCGGAGTATAGGCCCAC</sequence>
```

If acceptable in the context within which the sequence is being used, ambiguous nucleotide codes (M, R, W, S, Y, K, V, H, D, B, X, N) may be used in addition to primary nucleotide codes (A, C, G, T). For additional details on DNA sequence and IUPAC nucleic acid code usage within HML, refer to the DNA Sequence section appearing earlier in this document.

Change History for <sequence>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<SSO>

The *sso* element describes a sequence-specific oligonucleotide (SSO) test that was performed.

Attribute	Usage	Status	Description
ref-id	*HML 0.2+*	optional	Internal XML reference to a ***typing-test-names*** element within the document.
nmdp-refid	*HML 0.2+*	optional	ID of kit registered with NMDP. The cardinal sequence numbers of the registered probes in the kit determine the score order.
probe-name	*HML 0.2+*	optional	Fully qualified probe name (e.g. "L999.K1.V1.A9F-S11"). If this attribute is used, the scores attribute must contain exactly one score.
scores	*HML 0.2+*	required	Results of the SSO test, specified as one string (e.g. "118111100181").

One and only one of *ref-id*, *probe-name*, or *nmdp-refid* must be specified. The *ref-id*, *probe-name*, or *nmdp-refid* must specify a reference that resolves which SSO probes correspond to the *scores*.

Change History for <sso>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<SSP>

The *ssp* element describes a sequence-specific primer (SSP) test that was performed.

Attribute	Usage	Status	Description
ref-id	*HML 0.2+*	optional	Internal XML reference to a ***typing-test-names*** element within the document.
nmdp-refid	*HML 0.2+*	optional	ID of kit registered with NMDP. The cardinal sequence numbers of the registered SSPs in the kit determine the score order.
probe-name	*HML 0.2+*	optional	Fully qualified SSP name (e.g. "L999.K1.SSP12345"). If this attribute is used, the scores attribute must contain exactly one score.
scores	*HML 0.2+*	required	Results of the SSP test, specified as one string (e.g. "118111100181").

One and only one of *ref-id*, *probe-name*, or *nmdp-refid* must be specified. The *ref-id*, *probe-name*, or *nmdp-refid* must specify a reference that resolves which SSP probes correspond to the *scores*.

Change History for <ssp>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

<typing>

The *typing* element encapsulates a typing call (*interpretation* child element), combined with primary data (*sso*, *ssp*, and/or *sbt* child elements) which may have been used to determine that call.

Child Element	Usage	Status	Multiplicity	Notes
interpretation	*HML 0.3+*	optional	0 or 1	
sso	*HML 0.3+*	optional	0 or more	
ssp	*HML 0.3+*	optional	0 or more	
sbt	*HML 0.3+*	optional	0 or more	

Change History for <typing>

HML Version	Change Description
HML 0.3.3	Introduced *typing* element.

<typing-test-name>

The *typing-test-name* element specifies a test name.

Attribute	Usage	Status	Description
name	*HML 0.2+*	required	fully qualified test name (e.g. "L999.K1.V1.A9F-S11", "L999.K1.V1.SSP12345")

Change History for <typing-test-name>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

\<typing-test-names\>

The *typing-test-names* element specifies a list of test names internally referenced by an *sso* or *ssp* element.

Attribute	Usage	Status	Description
ref-id	*HML 0.2+*	required	XML ID reference unique within the document

Child Element	Usage	Status	Multiplicity	Notes
typing-test-name	*HML 0.2+*	required	1 or more	

Change History for \<typing-test-names\>

HML Version	Change Description
HML 0.2	Unchanged since *HML 0.2*.

References

1. Histoimmunogenetics Markup Language (HML)
http://www.bioinformatics.nmdp.org/HLA/HLA_Typing/HML/index.html

2. Maintaining updated DNA-based HLA assignments in the National Marrow Donor Program Bone Marrow Registry.
http://www.ncbi.nlm.nih.gov/pubmed/12361088

3. Nomenclature for Incompletely Specified Bases in Nucleic Acid Sequences
http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html

Appendix A: HML 0.3.3 DTD

```
<!--                                                                    -->
<!--  HML - Histo Immunogenetics Markup Language                        -->
<!--  DTD Version 0.3.3 (Draft)                                         -->
<!--  Written by: Jon Sorbie                                            -->
<!--  Modified by: Martin Maiers                                        -->
<!--                                                                    -->
<!--    COPYRIGHT (C) 2008 NATIONAL MARROW DONOR PROGRAM.               -->
<!--                       ALL RIGHTS RESERVED                          -->
<!--                                                                    -->
<!-- This document defines the HML markup language DTD used for reporting -->
<!-- primary typing data.                                               -->
<!--                                                                    -->
<!--  Version 0.3.3   Maiers  Make amplification-primer optional        -->
<!--
    hml is the root element of the document.  Any document level properties
    may be set with child property elements.
    Attributes:
    - version: The version of HML that this document should conform to.
-->
<!ELEMENT hml (property*, project+)>
<!ATTLIST hml
    version CDATA #REQUIRED
>


<!--
    The project element groups data into sections with a common destination
    or purpose, which is denoted by the "name" attribute.
    Attributes:
    - name: identifier for data's intent. (ex: "STAR", "SG09")
-->
<!ELEMENT project (reporting-center, sample+, typing-test-names*)>
<!ATTLIST project
    name  CDATA  #REQUIRED
>


<!--
    The reporting-center element specifies the origin of the data as a
    center code.
    Attributes:
    - code: reporting center code identifier
-->
<!ELEMENT reporting-center EMPTY>
<!ATTLIST reporting-center
    code  CDATA  #REQUIRED
>


<!--
    A "sample" element encloses the typing data pertaining to a particular
    sample. It may contain multiple typing elements.  One for each locus,
    for instance.
    Attributes:
    - id:         Identifier for the sample
    - center-code: center code of the sample's origin (donor center,
                   transplant center, etc.
-->
```

```
<!ELEMENT sample (property*, typing+)>
<!ATTLIST sample
    id           CDATA  #REQUIRED
    center-code  CDATA  #REQUIRED
>


<!--
    The typing element encapsulates a typing call (interpretation) combined
    with other primary data that may have been used to determine that call
    (sso, ssp, and/or sbt elements).
    Contains:
      1. at most one interpretation element
      2. as many primary data elements as needed
-->
<!ELEMENT typing (interpretation?, (sso | ssp | sbt)*)>


<!--
    interpretation element specifies the typing call at allele/code level.
    Usually contains one or two haploid elements for a particular locus, but
    may contain more if multiple loci are covered (example: two DRB1
    haploids + one DRB3 haploid).  As an alternative, it may contain genotype
    list(s) to represent the type(s).
    Attributes:
    - date: Date on which the typing was carried out, or on which the final
            call was determined.  Format can be either ISO-8601 or "YYYYMMDD".
-->
<!ELEMENT interpretation (haploid+ | genotype-list+)>
<!ATTLIST interpretation
    date  CDATA  #REQUIRED
>


<!--
    A haploid element specifies one-half of a full typing at a particular
    locus.
    Attributes:
    - locus:  Locus (ex: "A", "DRB1")
    - method: Typing method used (ex: "DNA", "SEROLOGY")
    - type:   Allele/code level type (ex: "0101", "01AB")
-->
<!ELEMENT haploid EMPTY>
<!ATTLIST haploid
    locus   CDATA  #REQUIRED
    method  CDATA  #REQUIRED
    type    CDATA  #REQUIRED
>


<!--
    A genotype-list represents a full unambiguous list of possibilities for
    the typing of a sample.  The values of the elements in this genotype-list
    (each allele element) should conform to the naming nomenclature used by the
    the source specified by the db attribute (if provided) or should be fully
    understood by the message recipient (if not provided).
    Attributes:
    - db:      Database or other source from which elements in this
               genotype-list are taken (ex: "HLADB")
    - version: A specific version of the aforementioned source (ex: "2.8.0")
-->
<!ELEMENT genotype-list (diploid-combination+)>
<!ATTLIST genotype-list
```

```
    db       CDATA  #IMPLIED
    version  CDATA  #IMPLIED
>


<!--
    A diploid-combination element is one possibility value in a genotype list.
    There may be either one or two locus-block child elements, depending on
    whether the data provided in this diploid-combination covers one or two
    chromosomes.
-->
<!ELEMENT diploid-combination (locus-block, locus-block?)>

<!--
    A locus-block element allows allele-list elements to be grouped together
    to mean one allele-list is a possibility if and only if all others are.
    This is useful, for example, in the case when listing DRB1 alleles
    next to the corresponding DRB3 alleles that are relevent in only some
    cases:
    <locus-block>
        <allele-list>
            <allele>DRB1*030101</allele>
            <allele>DRB1*0303</allele>
        </allele-list>
        <allele-list>
            <allele>DRB3*0203</allele>
        </allele-list>
    </locus-block>
-->
<!ELEMENT locus-block (allele-list+)>

<!--
    An allele list element is a representation of the list of allele
    possibilities for a type.  NMDP has historically used allele codes in
    combination with allele families to represent this.
    Example: A*01BC can now be represented as:
    <allele-list>
        <allele>A*0102</allele>
        <allele>A*0103</allele>
    </allele-list>
-->
<!ELEMENT allele-list (allele+)>

<!--
    An allele element specifies a single allele.  It should be given in
    LOCUS*NAME format and names must be at allele-level resolution.  The value
    should conform to the naming nomenclature used by the the source specified
    in the enclosing genotype-list element (if it was provided) or should be
    fully understood by the message recipient (if it was not).
    Example: <allele>A*010101</allele>
    Attributes:
    - present: Indicates presence of absence of this allele.  A value of "N"
               can be used to indicate that a particualar allele was tested for
               and found not to be a possibility.  A value of "U" (untested)
               indicates that the given allele was not tested for.  The default
               value is "Y"
-->
<!ELEMENT allele (#PCDATA)>
<!ATTLIST allele
    present  (Y|N|U)  "Y"
```

```
>


<!--
    Specifies an SSO test that was done.  Scores are required, along with a
    reference for resolving what SSO probes correspond to the list of scores.
    Only one of "ref-id", "probe-name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:      Internal XML reference to a typing-test-names element
                   contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                   numbers of the registered probes in the kit will determine
                   the score order.
    - probe-name: Fully qualified probe name.  (ex: "L999.K1.V1.A9F-S11")
                   If this attribute is used, the scores attribute must
                   contain exactly one score.
    - scores:      The results of the SSO test, specified as one string
                   (ex: "118111100181")
-->
<!ELEMENT sso EMPTY>
<!ATTLIST sso
    ref-id       IDREF  #IMPLIED
    nmdp-refid  CDATA  #IMPLIED
    probe-name  CDATA  #IMPLIED
    scores       CDATA  #REQUIRED
>


<!--
    Specifies an SSP test that was done.  Scores are required, along with a
    reference for resolving what SSPs correspond to the list of scores.  Only
    one of "ref-id", "name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:      Internal XML reference to a typing-test-names element
                   contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                   numbers of the registered SSPs in the kit will determine
                   the score order.
    - name:        Fully qualified SSP name.  (ex: "L999.K1.V1.SSP12345")
                   If this attribute is used, the scores attribute must
                   contain exactly one score.
    - scores:      The results of the SSP test, specified as one string
                   (ex: "118111100181")
-->
<!ELEMENT ssp EMPTY>
<!ATTLIST ssp
    ref-id       IDREF  #IMPLIED
    nmdp-refid  CDATA  #IMPLIED
    name         CDATA  #IMPLIED
    scores       CDATA  #REQUIRED
>


<!--
########## Begin SBT changes/additions ###################################
-->


<!--
    Specifies the details of an SBT test that was done.
-->
<!ELEMENT sbt (amplification-primer?, gssp?, sequence)>
```

```
<!--
    Amplification primer used.  Consists of the primer's registered name
    and the amplification score.  Score is usually '8', but may be '1' in which
    case no amplification happened and no sbt sequence is likely to have been
    obtained.
-->
<!ELEMENT amplification-primer (registered-name, score)>


<!--
    An identifying name that should be recognized by the message recipient.
-->
<!ELEMENT registered-name (#PCDATA)>


<!--
    Score - '8' or '1'
-->
<!ELEMENT score (#PCDATA)>


<!--
    Element to describe Group Specific Sequencing Primer (GSSP) used:
    - sequence if GSSP sequence is embedded within this document
    - registered-name if sequence has been registed with message recipient
-->
<!ELEMENT gssp (sequence | registered-name)?>


<!--
    DNA sequence consisting of primary nucleotides (A, C, G, T).
    Wildcard nucleotides (M, R, W, S, Y, K, V, H, D, B, X, N) may be used if
    they are acceptable in the context in which they appear.
    (ex: "<sequence>CCGGAGTATAGGCCCAC</sequence>")
-->
<!ELEMENT sequence (#PCDATA)>


<!--
########## End SBT changes/additions ################################
-->


<!--
    The typing-test-names element specifies a list of test names internally
    referenced by an sso element or an ssp element.  It wraps a list of
    "typing-test-name" elements.
    Attributes:
    - ref-id: XML ID reference unique in this document.
-->
<!ELEMENT typing-test-names (typing-test-name+)>
<!ATTLIST typing-test-names
    ref-id  ID  #REQUIRED
>


<!--
    A typing-test-name element specifies a test name contained in a
    referenced "typing-test-names" list.
    Attributes:
    - name: Fully qualified test name
            (ex: "L999.K1.V1.A9F-S11", "L999.K1.V1.SSP12345")
-->
<!ELEMENT typing-test-name EMPTY>
<!ATTLIST typing-test-name
    name  CDATA  #REQUIRED
```

```
>

<!--
    A property is a generic name-value pair
-->
<!ELEMENT property EMPTY>
<!ATTLIST property
    name   CDATA  #REQUIRED
    value  CDATA  #REQUIRED
>
```

Appendix B: HML 0.3 DTD

```
<!--                                                               -->
<!--  HML - Histo Immunogenetics Markup Language                   -->
<!--  DTD Version 0.3                                              -->
<!--  Written by: Jon Sorbie                                       -->
<!--                                                               -->
<!--    COPYRIGHT (C) 1994, 1995, 1997 NATIONAL MARROW DONOR PROGRAM.  -->
<!--                    ALL RIGHTS RESERVED                        -->
<!--                                                               -->
<!-- This document defines the HML markup language DTD used for reporting  -->
<!-- primary typing data.                                          -->
<!--                                                               -->

<!--
    hml is the root element of the document.  Any document level properties
    may be set with child property elements.
    Attributes:
    - version: The version of HML that this document should conform to.
-->
<!ELEMENT hml (property*, project+)>
<!ATTLIST hml
    version CDATA #REQUIRED
>

<!--
    The project element groups data into sections with a common destination
    or purpose, which is denoted by the "name" attribute.
    Attributes:
    - name: identifier for data's intent. (ex: "STAR", "SG09")
-->
<!ELEMENT project (reporting-center, sample+, typing-test-names*)>
<!ATTLIST project
    name  CDATA  #REQUIRED
>

<!--
    The reporting-center element specifies the origin of the data as a
    center code.
    Attributes:
    - code: reporting center code identifier
-->
<!ELEMENT reporting-center EMPTY>
<!ATTLIST reporting-center
    code  CDATA  #REQUIRED
>

<!--
    A "sample" element encloses the typing data pertaining to a particular
    sample. It may contain multiple typing elements.  One for each locus,
    for instance.
    Attributes:
    - id:        Identifier for the sample
    - center-code: center code of the sample's origin (donor center,
                 transplant center, etc.
-->
<!ELEMENT sample (property*, typing+)>
```

```
<!ATTLIST sample
    id           CDATA  #REQUIRED
    center-code  CDATA  #REQUIRED
>


<!--
    The typing element encapsulates a typing call (interpretation) combined
    with other primary data that may have been used to determine that call
    (sso, ssp, and/or sbt elements).
    Contains:
      1. at most one interpretation element
      2. as many primary data elements as needed
-->
<!ELEMENT typing (interpretation?, (sso | ssp | sbt)*)>


<!--
    interpretation element specifies the typing call at allele/code level.
    Usually contains one or two haploid elements for a particular locus, but
    may contain more if multiple loci are covered (example: two DRB1
    haploids + one DRB3 haploid).  As an alternative, it may contain genotype
    list(s) to represent the type(s).
    Attributes:
    - date: Date on which the typing was carried out, or on which the final
            call was determined.  Format can be either ISO-8601 or "YYYYMMDD".
-->
<!ELEMENT interpretation (haploid+ | genotype-list+)>
<!ATTLIST interpretation
    date  CDATA  #REQUIRED
>


<!--
    A haploid element specifies one-half of a full typing at a particular
    locus.
    Attributes:
    - locus:  Locus (ex: "A", "DRB1")
    - method: Typing method used (ex: "DNA", "SEROLOGY")
    - type:   Allele/code level type (ex: "0101", "01AB")
-->
<!ELEMENT haploid EMPTY>
<!ATTLIST haploid
    locus   CDATA  #REQUIRED
    method  CDATA  #REQUIRED
    type    CDATA  #REQUIRED
>


<!--
    A genotype-list represents a full unambiguous list of possibilities for
    the typing of a sample.  The values of the elements in this genotype-list
    (each allele element) should conform to the naming nomenclature used by the
    the source specified by the db attribute (if provided) or should be fully
    understood by the message recipient (if not provided).
    Attributes:
    - db:      Database or other source from which elements in this
               genotype-list are taken (ex: "HLADB")
    - version: A specific version of the aforementioned source (ex: "2.8.0")
-->
<!ELEMENT genotype-list (diploid-combination+)>
<!ATTLIST genotype-list
    db        CDATA  #IMPLIED
```

```
    version  CDATA  #IMPLIED
>


<!--
    A diploid-combination element is one possibility value in a genotype list.
    There may be either one or two locus-block child elements, depending on
    whether the data provided in this diploid-combination covers one or two
    chromosomes.
-->
<!ELEMENT diploid-combination (locus-block, locus-block?)>


<!--
    A locus-block element allows allele-list elements to be grouped together
    to mean one allele-list is a possibility if and only if all others are.
    This is useful, for example, in the case when listing DRB1 alleles
    next to the corresponding DRB3 alleles that are relevent in only some
    cases:
    <locus-block>
        <allele-list>
            <allele>DRB1*030101</allele>
            <allele>DRB1*0303</allele>
        </allele-list>
        <allele-list>
            <allele>DRB3*0203</allele>
        </allele-list>
    </locus-block>
-->
<!ELEMENT locus-block (allele-list+)>


<!--
    An allele list element is a representation of the list of allele
    possibilities for a type.  NMDP has historically used allele codes in
    combination with allele families to represent this.
    Example: A*01BC can now be represented as:
    <allele-list>
        <allele>A*0102</allele>
        <allele>A*0103</allele>
    </allele-list>
-->
<!ELEMENT allele-list (allele+)>


<!--
    An allele element specifies a single allele.  It should be given in
    LOCUS*NAME format and names must be at allele-level resolution.  The value
    should conform to the naming nomenclature used by the the source specified
    in the enclosing genotype-list element (if it was provided) or should be
    fully understood by the message recipient (if it was not).
    Example: <allele>A*010101</allele>
    Attributes:
    - present: Indicates presence of absence of this allele.  A value of "N"
               can be used to indicate that a particualar allele was tested for
               and found not to be a possibility.  A value of "U" (untested)
               indicates that the given allele was not tested for.  The default
               value is "Y"
-->
<!ELEMENT allele (#PCDATA)>
<!ATTLIST allele
    present  (Y|N|U)  "Y"
>
```

```
<!--
    Specifies an SSO test that was done.  Scores are required, along with a
    reference for resolving what SSO probes correspond to the list of scores.
    Only one of "ref-id", "probe-name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:     Internal XML reference to a typing-test-names element
                  contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                  numbers of the registered probes in the kit will determine
                  the score order.
    - probe-name: Fully qualified probe name.  (ex: "L999.K1.V1.A9F-S11")
                  If this attribute is used, the scores attribute must
                  contain exactly one score.
    - scores:     The results of the SSO test, specified as one string
                  (ex: "118111100181")
-->
<!ELEMENT sso EMPTY>
<!ATTLIST sso
    ref-id       IDREF  #IMPLIED
    nmdp-refid  CDATA  #IMPLIED
    probe-name  CDATA  #IMPLIED
    scores      CDATA  #REQUIRED
>


<!--
    Specifies an SSP test that was done.  Scores are required, along with a
    reference for resolving what SSPs correspond to the list of scores.  Only
    one of "ref-id", "name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:     Internal XML reference to a typing-test-names element
                  contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                  numbers of the registered SSPs in the kit will determine
                  the score order.
    - name:       Fully qualified SSP name.  (ex: "L999.K1.V1.SSP12345")
                  If this attribute is used, the scores attribute must
                  contain exactly one score.
    - scores:     The results of the SSP test, specified as one string
                  (ex: "118111100181")
-->
<!ELEMENT ssp EMPTY>
<!ATTLIST ssp
    ref-id       IDREF  #IMPLIED
    nmdp-refid  CDATA  #IMPLIED
    name        CDATA  #IMPLIED
    scores      CDATA  #REQUIRED
>


<!--
    Specifies an SBT test that was done.  The result of the test (the DNA
    sequence) is given in the child "sequence" element.
    Attributes:
    - name:  Fully qualified name of the SBT test (ex: "L999.K1.V1.SBT54321")
    - score: Result of primer amplification used for this test.  This is
             usually '8', but may be '1' in which case no amplification
             happened and no sequence is likely to have been obtained.
-->
<!ELEMENT sbt (sequence)>
```

```
<!ATTLIST sbt
    name    CDATA   #REQUIRED
    score   CDATA   #REQUIRED
>


<!--
    The sequence element contains the DNA sequence obtained from an SBT
    test.  Primary (A, C, G, T) and wildcard nucleotides
    (M, R, W, S, Y, K, V, H, D, B, X, N) may be used
    (ex: "<sequence>CCGGAGTATTGGGACCAGGAGACACGGAATATGAAGGCCCAC</sequence>")
-->
<!ELEMENT sequence (#PCDATA)>


<!--
    The typing-test-names element specifies a list of test names internally
    referenced by an sso element or an ssp element.  It wraps a list of
    "typing-test-name" elements.
    Attributes:
    - ref-id: XML ID reference unique in this document.
-->
<!ELEMENT typing-test-names (typing-test-name+)>
<!ATTLIST typing-test-names
    ref-id  ID  #REQUIRED
>


<!--
    A typing-test-name element specifies a test name contained in a
    referenced "typing-test-names" list.
    Attributes:
    - name: Fully qualified test name
            (ex: "L999.K1.V1.A9F-S11", "L999.K1.V1.SSP12345")
-->
<!ELEMENT typing-test-name EMPTY>
<!ATTLIST typing-test-name
    name  CDATA  #REQUIRED
>


<!--
    A property is a generic name-value pair
-->
<!ELEMENT property EMPTY>
<!ATTLIST property
    name    CDATA  #REQUIRED
    value   CDATA  #REQUIRED
>
```

Appendix C: HML 0.2 DTD

```
<!--                                                                    -->
<!--  HML – Histo Immunogenetics Markup Language                        -->
<!--  DTD Version 0.2 Alpha                                             -->
<!--  Written by: Jon Sorbie                                            -->
<!--                                                                    -->
<!--    COPYRIGHT (C) 1994, 1995, 1997 NATIONAL MARROW DONOR PROGRAM.   -->
<!--                       ALL RIGHTS RESERVED                          -->
<!--                                                                    -->
<!-- This document defines the HML markup language DTD used for reporting -->
<!-- primary typing data.                                               -->
<!--                                                                    -->

<!--
    hml is the root element of the document.
    Attributes:
    - version: The version of HML that this document should conform to.
-->
<!ELEMENT hml (project+)>
<!ATTLIST hml
    version CDATA #REQUIRED
>

<!--
    The project element groups data into sections with a common destination
    or purpose, which is denoted by the "name" attribute.
    Attributes:
    - name: identifier for data's intent. (ex: "STAR", "SG09")
-->
<!ELEMENT project (reporting-center, sample+, typing-test-names*)>
<!ATTLIST project
    name  CDATA  #REQUIRED
>

<!--
    The reporting-center element specifies the origin of the data as a
    center code.
    Attributes:
    - code: reporting center code identifier
-->
<!ELEMENT reporting-center EMPTY>
<!ATTLIST reporting-center
    code  CDATA  #REQUIRED
>

<!--
    A "sample" element encloses the typing data pertaining to a particular
    sample. It may contain multiple hla-typing elements.  One for each locus,
    for instance.
    Attributes:
    - nmdp-id:     NMDP ID in non-hyphenated format.  Need not be zero-padded.
    - center-code: center code of the sample's origin (donor center,
                   transplant center, etc.
-->
<!ELEMENT sample (hla-typing+)>
<!ATTLIST sample
```

```
        id          CDATA  #REQUIRED
        center-code CDATA  #REQUIRED
>


<!-- Entity reference for one of the subelements of hla-typing. -->
<!ENTITY % hla-typing-child "sso | ssp | sbt">


<!--
    The hla-typing element encapsulates a typing call combined with other
    primary data that may have been used to determine that call (sso, ssp,
    and/or sbt elements).
    Contains:
      1. at most one interpretation element
      2. as many hla-typing-child elements as needed
    These may appear in any order, but at least one child must be present.
-->
<!ELEMENT hla-typing ((%hla-typing-child;)*,
                      (interpretation | %hla-typing-child;),
                      (%hla-typing-child;)*)>


<!--
    interpretation element specifies the typing call at allele/code level.
    Usually contains one or two haploid elements for a particular locus, but
    may contain more if multiple loci are covered (example: two DRB1
    haploids + one DRB3 haploid).  As an alternative, it may contain one or
    more allele lists to represent the type(s).
    Attributes:
    - date: Date on which the typing was carried out, or on which the final
            call was determined.  Required to be in "YYYYMMDD" format
            (ex: "20030101")
-->
<!ELEMENT interpretation (haploid+ | allele-list+)>
<!ATTLIST interpretation
    date  CDATA  #REQUIRED
>


<!--
    A haploid element specifies one-half of a full typing at a particular
    locus.
    Attributes:
    - locus:  Locus (ex: "A", "DRB1")
    - method: Typing method used (ex: "DNA", "SEROLOGY")
    - type:   Allele/code level type (ex: "0101", "01AB")
-->
<!ELEMENT haploid EMPTY>
<!ATTLIST haploid
    locus   CDATA  #REQUIRED
    method  CDATA  #REQUIRED
    type    CDATA  #REQUIRED
>


<!--
    An allele list is a representation of the list of possibilities for a
    type.
    Contains one or more allele elements
-->
<!ELEMENT allele-list (allele+)>


<!--
```

```
    An allele element is simply the designation for a single allele.
    Attributes:
    - source: Identifier (namespace) for the name attribute.  This can be
              thought of as the HLA sequence database release version
              (ex: "HLADB-2.0.0")
    - name:   Identifier for the allele (The allele name) (ex: "A*010101")
-->
<!ELEMENT allele EMPTY>
<!ATTLIST allele
    source CDATA #REQUIRED
    name   CDATA #REQUIRED
>


<!--
    Specifies an SSO test that was done.  Date and scores are required, along
    with a reference for resolving what SSO probes correspond to the list of
    scores.  Only one of "ref-id", "probe-name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:     Internal XML reference to a typing-test-names element
                  contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                  numbers of the registered probes in the kit will determine
                  the score order.
    - probe-name: Fully qualified probe name.  (ex: "L999.K1.V1.A9F-S11")
                  If this attribute is used, the scores attribute must
                  contain exactly one score.
    - scores:     The results of the SSO test, specified as one string
                  (ex: "118111100181")
-->
<!ELEMENT sso EMPTY>
<!ATTLIST sso
    ref-id       IDREF  #IMPLIED
    nmdp-refid   CDATA  #IMPLIED
    probe-name   CDATA  #IMPLIED
    scores       CDATA  #REQUIRED
>


<!--
    Specifies an SSP test that was done.  Date and scores are required, along
    with a reference for resolving what SSPs correspond to the list of
    scores.  Only one of "ref-id", "name", or "nmdp-refid" may be given.
    Attributes:
    - ref-id:     Internal XML reference to a typing-test-names element
                  contained in this document.
    - nmdp-refid: ID of kit registered with NMDP.  The cardinal sequence
                  numbers of the registered SSPs in the kit will determine
                  the score order.
    - name:       Fully qualified SSP name.  (ex: "L999.K1.V1.SSP12345")
                  If this attribute is used, the scores attribute must
                  contain exactly one score.
    - scores:     The results of the SSP test, specified as one string
                  (ex: "118111100181")
-->
<!ELEMENT ssp EMPTY>
<!ATTLIST ssp
    ref-id       IDREF  #IMPLIED
    nmdp-refid   CDATA  #IMPLIED
    name         CDATA  #IMPLIED
    scores       CDATA  #REQUIRED
```

```
>


<!--
    Specifies an SBT test that was done.  The result of the test (the DNA
    sequence) is given in the child "sequence" element.
    Attributes:
    - name:  Fully qualified name of the SBT test (ex: "L999.K1.V1.SBT54321")
    - score: Result of primer amplification used for this test.  This is
             usually '8', but may be '1' in which case no amplification
             happened and no sequence is likely to have been obtained.
-->
<!ELEMENT sbt (sequence)>
<!ATTLIST sbt
    name   CDATA  #REQUIRED
    score  CDATA  #REQUIRED
>


<!--
    The sequence element contains the DNA sequence obtained from an SBT
    test.  Primary (A, C, G, T) and wildcard nucleotides
    (M, R, W, S, Y, K, V, H, D, B, X, N) may be used
    (ex: "<sequence>CCGGAGTATTGGGACCAGGAGACACGGAATATGAAGGCCCAC</sequence>")
-->
<!ELEMENT sequence (#PCDATA)>


<!--
    The typing-test-names element specifies a list of test names internally
    referenced by an sso element or an ssp element.  It wraps a list of
    "typing-test-name" elements.
    Attributes:
    - ref-id: XML ID reference unique in this document.
-->
<!ELEMENT typing-test-names (typing-test-name+)>
<!ATTLIST typing-test-names
    ref-id  ID  #REQUIRED
>


<!--
    A typing-test-name element specifies a test name contained in a
    referenced "typing-test-names" list.
    Attributes:
    - name: Fully qualified test name
            (ex: "L999.K1.V1.A9F-S11", "L999.K1.V1.SSP12345")
-->
<!ELEMENT typing-test-name EMPTY>
<!ATTLIST typing-test-name
    name  CDATA  #REQUIRED
>
```