**Hi-C file layout**

**Header**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Type** | **Value** |
| Magic | Hi-C magic string | String | HIC |
| Version | Version number | int | 8 |
| mIdxPos | File position of master index | long |  |
| genomeId | Genome identifier (e.g. hg19, mm9, etc) | String |  |
| nAttributes | Number of key-value pair attributes | int |  |
| |  |  |  |  | | --- | --- | --- | --- | | *List of key-value pair attributes (n = nAttributes)* | | | | | key |  | String |  | | value |  | String |  | | | | |
| nChrs | Number of chromosomes |  |  |
| |  |  |  |  | | --- | --- | --- | --- | | *List of chromosome lengths (n = nChrs)* | | | | | chrName | Chromsome name | String |  | | chrLength | Chromosome length | int |  | | | | |
| nBpResolutions | Number of base pair resolutions | int |  |
| |  |  |  |  | | --- | --- | --- | --- | | *List of bin sizes for bp resolution levels (n = nBpResolutions)* | | | | | resBP | Bin size in base pairs | int |  | | | | |
| nFragResolutions | Number of fragment resolutions | int |  |
| |  |  |  |  | | --- | --- | --- | --- | | *List of bin sizes for frag resolution levels (n = nFragResolutions)* | | | | | resFrag | Bin size in fragment units (1, 2, 5, etc) | int |  | | | | |
| |  |  |  |  | | --- | --- | --- | --- | | *List of fragment sites per chromosome, in same order as chromosome list above (n = nChrs). This section absent if nFragResolutions == 0.)* | | | | | nSites | Number of sites for this chromosome | int |  | | |  |  |  |  | | --- | --- | --- | --- | | *List of sites (n = nSites)* | | | | | site | Site position in base pairs | int |  | | | | | | | | |

**Body**

|  |  |  |  |
| --- | --- | --- | --- |
| **Field** | **Description** | **Type** | **Value** |
| *The section below is repeated for each chr-chr combination* | | | |
| chr1Idx | Chr 1 index | int |  |
| chr2Idx | Chr 2 index | int |  |
| nResolutions | Total number of resolutions, including base pair and fragment resolutions. | int |  |
| |  |  |  |  | | --- | --- | --- | --- | | *Resolution headers (n = nResolutions)* | | | | | unit | Distance unit, base-pairs or fragments | String | FRAG : BP | | resIdx | Index number for this resolution level. | int |  | | sumCounts | Sum of all counts (or scores) across all bins | float |  | | occupiedCellCount | Total count of cells that are occupied (currently unused) | float |  | | stdDev | Standard deviation (currently unused) | float |  | | percent95 | Estimate of 95th percentile of counts among occupied bins (currently unused) | float |  | | binSize | The bin size in base-pairs or fragments | int |  | | blockBinCount | Number of bins divided by number of columns (around 1000) | int |  | | blockColumnCount | Number of bins divided by 1000 | int |  | | blockCount | Number of blocks | int |  | | |  |  |  |  | | --- | --- | --- | --- | | *Block index (n = nResolutions)* | | | | | blockID | Numeric id for block | int |  | | blockPositiion | File position of block | long |  | | blockSize | Size of block in bytes | int |  | | | | | | | | |
| |  | | --- | | *Resolution level data (n = nResolutions)* | | |  |  |  |  | | --- | --- | --- | --- | | *Block data (n = blockCount)* | | | | | cellCount | Number of cells in this block (sparse matrix) | int |  | | |  |  |  |  | | --- | --- | --- | --- | | *Cell data (n = cellCount)* | | | | | binX | X axis index | int |  | | binY | Y axis index | long |  | | value | Value (counts or score) | float |  | | | | | | | | | |

**Footer**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Field** | **Description** | **Type** | | **Value** |
| nBytesV5 | Number of bytes for the “v5” footer, that is everything up to the normalized expected vectors | int | |  |
| *Master index for “Matrix” records* | | | | |
| nEntries | Number of index entries | int | |  |
| |  |  |  |  | | --- | --- | --- | --- | | *List of index entries (n = nEntries)* | | | | | key | Key, constructed as <chr1>\_<chr2> | String |  | | position | Position of start of matrix record in bytes | long |  | | size | Size of matrix record in bytes | int |  | | | | | |
| *Expected value vectors* | | | | |
| nExpectedValueVectors | Number of expected value vectors to follow. These are expected values from the non-normalized observed matrix. | int |  | |
| |  |  |  |  | | --- | --- | --- | --- | | *List of expected value vectors (n = nExpectedValueVectors)* | | | | | binSize | Bin (grid) size for this calculation | int |  | | unit | Bin units either FRAG or BP. | String | FRAG : BP | | nValues | Size of the vector | int |  | | |  |  |  |  | | --- | --- | --- | --- | | *List of expected values (n = nValues)* | | | | | value | Expected value | double |  | | | | | | nChrScaleFactors | | | | | |  |  |  |  | | --- | --- | --- | --- | | *List of normalization factors (n =* nChrScaleFactors*)* | | | | | chrIndex | Chromosome index | int |  | | chrScaleFactor | Chromosome scale factor | double |  | | | | | | | | | |
|  | | | | |
| nNormExpectedValueVectors | Number of normalized expected value vectors to follow | int |  | |
| |  |  |  |  | | --- | --- | --- | --- | |  | | | | | Type | Indicates type of normalization | String | VC:KR | | binSize | Bin (grid) size for this calculation | int |  | | unit | Bin units either FRAG or BP. | String | FRAG : BP | | nValues | Size of the vector | int |  | | |  |  |  |  | | --- | --- | --- | --- | | *List of expected values (n = nValues)* | | | | | value | Expected value | double |  | | | | | | nChrScaleFactors | | | | | |  |  |  |  | | --- | --- | --- | --- | | *List of normalization factors (n =* nChrScaleFactors*)* | | | | | chrIndex | Chromosome index | int |  | | chrScaleFactor | Chromosome scale factor | double |  | | | | | | | | | |
| *Normalization vector index* | | | | |
| nNormVectors | Number of normalization vectors | int |  | |
| |  |  |  |  | | --- | --- | --- | --- | | *List of normalization vectors (n= nNormalizationVectors)* | | | | | type | Indicates type of normalization | String | VC:KR | | chrIdx | Chromosome index | int |  | | unit | Bin units either FRAG or BP. | int | FRAG : BP | | binSize | Resolution | int |  | | Position | File position of value array | long |  | | nBytes | Size in bytes of value array | int |  | |  | | | | | | | | |
| *Normalization vector arrays (repeat for each entry above)* | | | | |
| nValues | Number of values in array | int |  | |
| |  |  |  |  | | --- | --- | --- | --- | | *Normalization vector values (n= nValues)* | | | | | value |  | double |  | | | | | |
| *Attribute dictionary* | | | | |
| nAttributes | Number of attributes | int |  | |
| |  |  |  |  | | --- | --- | --- | --- | | *List of key-value pair attributes (n = nAttributes)* | | | | | key |  | String |  | | value |  | String |  | | | | | |