**Project:**  Genome Notepad

**Version:**  1.0

**Document:** Software Specification (Short Form)

**Company:** Metalcom (PTY)

**Author:** Rohin Gosling

**Description:**

DNA / RNA sequence editor.

**Features:**

* File Formats
  + Open and Save
    - Proprietary genetic sequence file format.
      * \*.SEQ

Basic single sequence file format used by Genome Notepad.

* + - * \*.SEQX

Advanced multi sequence file format used by Genome Studio.

* + Import and Export
    - Plain sequence format.
    - EMBL
    - FASTA
    - GCG format
    - GCG-RSF (Rich Sequence Format)
    - Genbank Format
    - IG Format
    - Genomatics annotation syntax
* File Size
  + Open and edit files of any size.
  + Open and edit files whose disk size is larger than available RAM.
* Notation
  + RNA
  + DNA
  + Ambiscript
* Sequencing
  + Support both 3’ and 5’ directions.
  + Support both sense and anti-sense strands.
  + Support multilayer sequences

**Components:**

* GUI
  + HEX editor inspired genetic sequence editor.
  + Text editor.
  + Layered sequence map.
  + Tree view sequence browser.
  + Properties window.
  + Linear and circular graphic map diagrams.

**Genome Notepad**

**File Header**

* File ID String
* Major Version
* Minor Version
* Project Name
* Project Description
* String Table Offset
* Base Group Offset
* Base Table Offset
* Codon Group Offset
* Codon Table Offset
* Sequence Allocation Table Offset
* Segment Type Table Offset
* Segment Allocation Table Offset
* Annotation Table Offset
* Maximum String Length
* Base Group Length
* Base Record Length
* Codon Group Length
* Codon Record Length
* Sequence Allocation Record Length
* Segment Type Record Length
* Segment Allocation Record Length
* Annotation Record Length
* String Count
* Base Count
* Codon Count
* Segment Type Count
* Highlight Count
* Sequence Count
* Segment Count

**String Record**

* String ID
* String Length
* String Data

**Base Record**

* Base ID
* Base Name
* Base Description
* Polarity ( Non-Polar, Polar )
* pH ( Basic, Acidic )

**Codon Record**

* Codon ID
* Codon Name
* Codon Description
* Codon Width

**Segment Type Record**

* Segment Type ID
* Segment Type Name
* Segment Type Description

Note:

Default segment types include:

* + Exon
  + Intron
  + Promoter
  + Enhancer
  + Operator
  + Repeat
  + Module
  + S/MAR
  + mRNA Cap
  + mRNA 5’ UTR
  + mRNA CDS (Coding Sequence)
  + mRNA 3’ UTR
  + mRNA Poly-A Tail

**Sequence Allocation Record**

* Sequence Allocation ID
* Sequence Allocation Name
* Sequence Allocation Description
* Topology (Linear, Circular)
* Origin

**Segment Allocation Record**

* Segment ID
* Parent Sequence ID
* Segment Name
* Segment Description
* Segment Type ID
* Codon ID
* Segment Direction (3’, 5’)
* \* (Sense, Anti-Sense)
* Segment Start Address
* Segment End Address

**Annotation Record**

* Annotation ID
* Annotation Name
* Annotation Description
* Start Address
* End Address
* Foreground Color
* Background Color
* Shape ( Box, 3’ Arrow, 5’ Arrow )