

IGV(Integrative Genomics Viewer)使用指南

一、软件说明

IGV 是一款针对高通量测序数据进行可视化的专业软件。可以帮助睿锶用户对分析结果中的相关文件在可视化的情况进行浏览分析。下载地址：

WINDOWS 系统：https://data.broadinstitute.org/igv/projects/downloads/2.19/IGV_Win_2.19.2-WithJava-installer.exe

MAC 系统（apple 芯片）：https://data.broadinstitute.org/igv/projects/downloads/2.19/IGV_MacApp_2.19.2_WithJava.zip

MAC 系统（inter 芯片）：https://data.broadinstitute.org/igv/projects/downloads/2.19/IGV_MacAppIntel_2.19.2_WithJava.zip

二、数据导入

IGV 可视化时，文件需要和比对时使用的参考基因组对应，您可以通过报告中的链接下载对应的参考基因组 fa 文件以及基因注释 gff 文件。

1. 导入本地基因组文件：

Genomes→load genomes from file→选择本地 fa/fastq 文件

2. 导入本地基因注释文件

File→load from files→选择本地 gff/gtf 文件

3. 导入 bam 文件

File→load from file→选择报告中的 bam(一般位于 2.map 中，同一文件夹下必须同时存在同名的.bam.bai 索引文件)

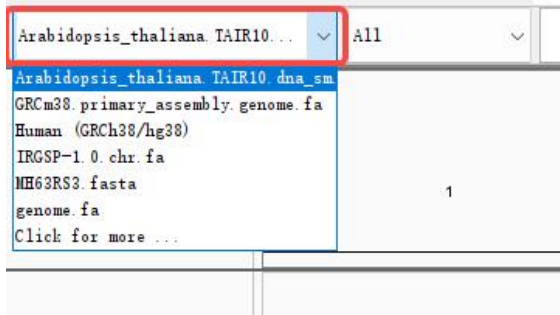
4. 导入其他格式文件

IGV 可以对多种格式文件进行可视化，bw 格式文件是简化后的 bam 文件，bed/narrowPeak/broadPeak 是前三列为基因组坐标的文件，narrowPeak/broadPeak 是 call peak 得到的 peak 信息文件。

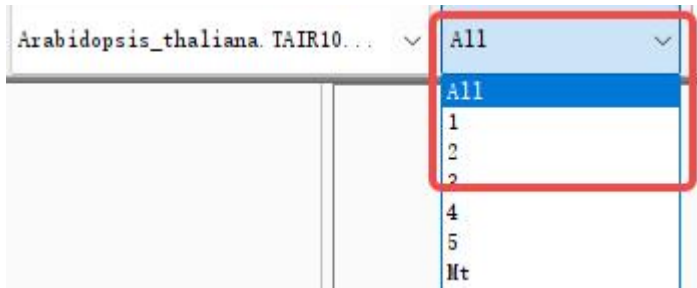
File→load from file→选择报告中的 bw/bed/narrowPeak/broadPeak 文件

三、软件界面说明

工具栏：提供一些常用的功能的按钮：



: 基因组下拉菜单，用于选择基因组。



: 染色体下拉菜单，用于选择染色体。



: 回到初始的全基因组界面。



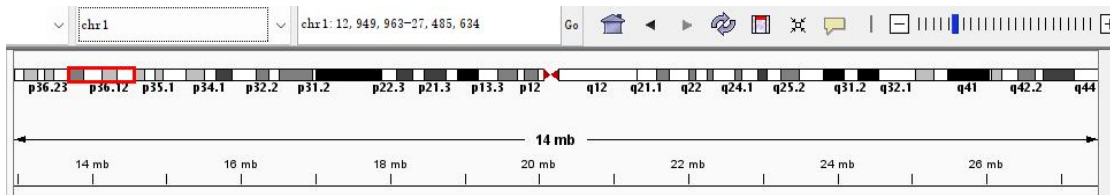
: 前进或后退一个画面。



选择一个区域并将其存为感兴趣的区域。



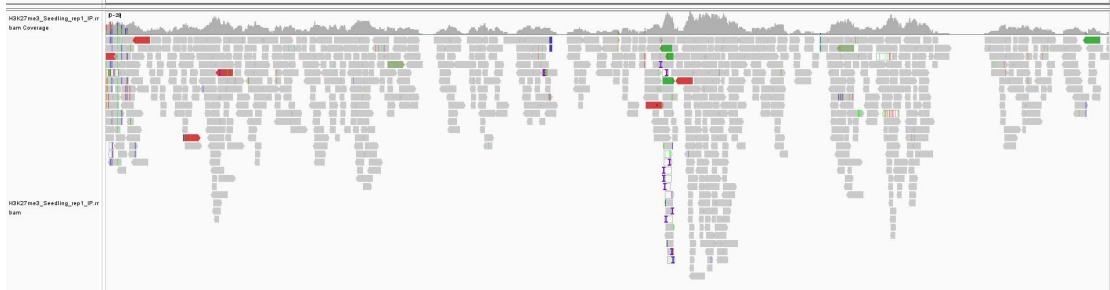
: 放大或缩小视野区域。(zoom in and zoom out)



染色体图：如果有对应的染色体核型图片，会在该区域展示，红色方框代表 当

前视野所展示的区域在染色体上的对应位置。

标尺：染色体图下方反应当前视野的宽度(图中是 14mb)，通过工具栏中的滑动按钮能改变此区域的大小，在该区域选中然后向左或向右拖动能移动当前的视野，结合放大缩小功能就能对文件任意部分进行查看。



比对结果展示区：当导入 bam 文件后，不同样品的数据会在该区域展示，每个 bam 文件会出现两个组成部分(2 个 track)，一个是覆盖度栏 (Coverage Track)，一个是比对结果栏 (reads 所在的栏)。



序列及注释文件展示区：基因组的序列、gtf 文件、bed 文件等导入后会在这个区域进行展示。

四、 软件核心用法说明

1. 保存目标区域

在浏览过程中，我们可以对特殊的位点进行一个存储，以便于之后快速找到这些区域，以及对不同的区域进行比较。我们可以通过工具栏中存储感兴趣的区域按钮进行实现。

点击保存区域图标，然后在视野中点击鼠标选择特定区域，该区域就被存储为一个感兴趣的区域。点击主菜单中的 Regions→Regions Navigator，我们可以进入一个关于感兴趣区域的列表，在列表中我们可以方便的对区域进行编辑。如下图，我们存储了两个感兴趣区域，通过 add，我们能将当前视野所在区域存储为一个感兴趣区域，通过在列表中选中区域并点击 remove，我们能删除区域。通过点击 view，视野会转移到之前所存储的感兴趣区域。

2. 设置中心线帮助浏览

放大缩小视野及移动视野的功能主要通过工具栏中的滑动按钮和标尺区域实现，具体的用法已经在界面说明中进行介绍，在这里我们介绍如果在放大过程中始终让目标区域处在视野正中央的位置，而不会随着放大走出视野。这个技巧对于浏览大型基因组时很有用。这项功能主要通过中心线(center line)实现，如图中红框所示。我们可以把想看的区域移动到中心线，然后点击放大，这样想看的区域就不会向左或向右移动脱离视野。

设置中心线：主菜单 View→Preferences→alignments→show center line

3. 查看序列信息

在导入基因组后，我们能在序列及注释文件展示区(界面介绍中的区域 5)查看序列及 **gtf** 注释的外显子信息。如下图 (基因序列只有在当前视野足够小时显示):

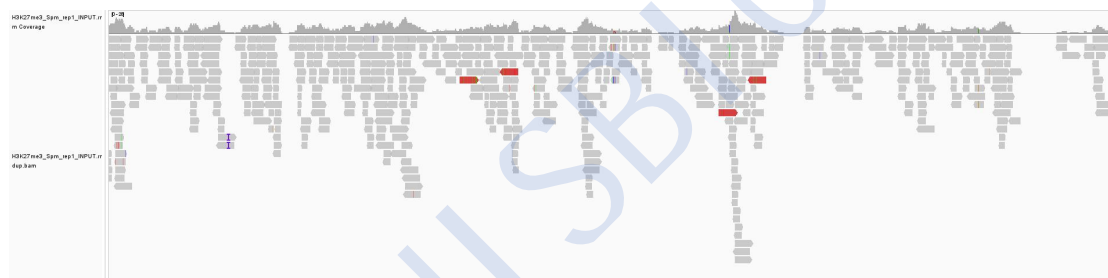
点击基因序列，我们能查看其对应的翻译后的氨基酸，三列氨基酸序列对应了三种 **reading frame** 下的氨基酸序列(如下图)。序列的箭头指向右表示当前显示的是正链，右键点击序列，选择 **Flip strand** 能改变成查看互补链信息。在 **Translation Table** 中能选择物种所对应的氨基酸密码子对应表。

4. 通过 bam 文件查看比对情况

在导入 **bam** 文件后，我们能在比对结果展示区(界面介绍中的区域 4)查看比对相关的信息。

覆盖度栏(Coverage Track): 主要用于直接展示 **reads** 的丰度，高度代表比对到此处的测序片段(**read**) 数目，当比对到参考基因组上的测序片段中有超过 20%与参考基因组不同时,就 用不同的颜色标注:红色-T,蓝色-C,绿色-A,橙色-G。否则就用灰色 标注。如下图所示，在更高的分辨率下,可以查看每个碱基下覆盖的各种碱基的数目及比例。此外，点击右键可选择不同的展示形式的选项。

Reads 栏(Reads Track): 主要用于直接展示 **reads** 的比对情况，右键点击 **reads track**，选择 **Expanded**，**Show mismatched bases**，能得到类似下图的展示结果。



图中实心灰代表比对质量比较高的测序片段,空心灰代表比对到此处的测序片段也可以比对到其他位点。高分辨率下,可以精确到每个位点的碱基类型:当比对序列上与参考基因组相同的超过 80%时,用灰色表示;否则用红色-T,蓝色-C,绿色-A,橙色-G,如上图中一些蓝色、绿色、红色及橙色的竖线。图中的黑点代表有插入或缺失的发生,放大后能查看细节,紫色代表插入,黑色横线代表缺失,鼠标选择后查看细节。

5. 保存图片

点击 **File**→**Save PNG Image** 或者 **Save SVG Image** 将当前视图保存为 **png** 图片或者 **svg** 矢量图。