

Here is a review of your bladder.Rmd R Notebook, focusing on finding bugs and suggesting improvements for speed and reliability:

---

## Bugs and Issues

---

### 1. Repeated Installation of Packages (Slows Execution)

- Every time the notebook runs, it attempts to install all required packages, even if they are already present. This is slow and unnecessary.

**Fix:**

Wrap each install in `if (!require(...)) install.packages(...)` for CRAN, and `if (!requireNamespace(...)) BiocManager::install(...)` for Bioconductor.

---

### 2. Loading Data: `lihc_res` vs. BLCA

- The code uses `lihc_res` as the variable name for results from a TCGA BLCA (bladder cancer) query. The prefix “lihc” usually refers to liver cancer in TCGA. This is confusing and may be a copy-paste artifact.

**Fix:**

Rename all instances of `lihc_res` to something like `blca_res` for clarity.

---

### 3. Use of `is_empty()` Without Loading `rlang`

- `is_empty()` is not a base R function; it's from the `rlang` or `purrr` package, but neither is loaded. This will cause an error unless the user has it loaded globally.

**Fix:**

Add `library(rlang)` or check for length/NULL with base R only.

---

### 4. Possible Bug in Heatmap Generation

- In the heatmap section, this line may cause an error:

```
t(d_mat[,limma_res$(topGenes$)gene_id])
```

The column `gene_id` may not exist or match the matrix columns if the names are not aligned.

**Fix:**

Double-check that `limma_res$(topGenes$)gene_id` matches the column names of `d_mat`. If using gene names, ensure they're mapped correctly.

## 5. Loading and Saving Workspaces

- The workflow heavily depends on loading and saving `.RData` and `.RDS` files. If the structure or objects change, loading old files can result in errors or inconsistent data.

**Fix:**

Add version checks or defensive programming when loading data. Always check that loaded objects are as expected.

## 6. VennDiagram: Argument Typo

- In both Venn diagrams, you use `fil=` instead of `fill=`, which will not work.

**Fix:**

Change `fil` to `fill` in the `venn.diagram` calls.

## 7. PCA Plot: Hardcoded Group Names

- You use:

```
\(targets[, condition_variable])
```

group = factor(voomObj ``` targets` may not  
but `voomObj\`  
contain the expected column if the pipeline changes.

**Fix:**

Add checks that the column exists, or handle errors gracefully.

---

## 8. Deprecated Functions

- Use of `write_rds` (should be `write_rds` from `readr`, but you use `readr::write_rds`). Make sure the right package is loaded.
- 

## 9. Inconsistent Use of `rowData`, `colData`, and Direct Slots

- Sometimes you use `@colData`, sometimes `colData()`, sometimes direct slot access. This is brittle.

**Fix:**

Prefer accessor functions ( `colData()`, `rowData()` ) for S4 objects for better compatibility.

---

## Suggestions for Speeding Up Execution

---

### A. Avoid Repeated Data Downloads

- Only download TCGA data if not already present. You already check for saved workspaces—consider also checking for the raw downloaded files.

### B. Use More Cores for WGCNA

- You set `enableWGCNAThreads(2)`. If your machine supports more, consider increasing this for speed.

### C. Parallelization

- For any for-loops or apply-family calls in downstream analyses (e.g., calculating module statistics), use `doParallel` or `foreach` for parallelization.

## D. Reduce Data Size Early

- Filter low-expressed genes as early as possible (before normalization) to reduce matrix sizes for downstream steps.

## E. Save Intermediate Files

- Save intermediate results (e.g., normalized expression matrices, DEG results) to disk and reload if available to avoid recomputation.

## F. Conditional Plotting

- Only generate plots if needed (e.g., add a flag for plotting). Plotting can be slow, especially for heatmaps.

---

## Miscellaneous Improvements

- Move all library calls to one chunk at the top.
  - Use `message()` or `cat()` to print progress updates.
  - Wrap long-running code in `suppressMessages()` or `suppressWarnings()` to keep output clean.
  - Clean up temporary variables with `rm()` and `gc()` to free memory, especially before WGCNA.
-

## Summary Table

Bug/Issue (Line/Section)	Quick Fix Suggestion
Unconditional package installs	Wrap install in <code>if (!require())</code>
Misnamed variables ( <code>lihc_res</code> )	Use <code>blca_res</code> or similar
Use of <code>is_empty()</code> w/o import	Use <code>length(x) == 0</code> or load <code>rlang</code>
Heatmap columns may not match	Ensure column names match between objects
VennDiagram <code>fil=</code> typo	Change <code>fil=</code> to <code>fill=</code>
Inconsistent slot access	Use accessor functions
Possible old workspace incompatibility	Add version checks/check object structure

Let me know if you want a revised file or specific rewritten code chunks!