Requirements Document: Gene Explorer - Drug Repurposer App

To: PoliHack Participants
From: QIAGEN Digital Insights
Date: April, 2025

Project Overview

We're excited to challenge you to build the **Gene Explorer - Drug Repurposer App**, a mobile or desktop application sponsored by QIAGEN Digital Insights. This app will allow users to explore a gene of interest, visualize its biological context, and propose existing drugs that could be repurposed to target it or related genes. Your app will combine bioinformatics exploration with a practical twist: drug repurposing, a cutting-edge approach to accelerate therapeutic discovery.

Objective

Create an interactive tool that:

1. Provides an overview of a user-specified gene.
2. Displays a gene-gene interaction network, including similar genes.
3. Identifies drugs associated with the gene of interest and its network, suggesting potential repurposing opportunities.

Functional Requirements

1. **Gene Input and Overview**
    - *Input*: Users enter a gene name (e.g., "BRCA1", "TP53") via a text field.

      Very rare genetic diseases may benefit from drug repurposing, so we could use some genes that are associated with a rare condition, and check whether an existing drug may target the same genes. You can use the following resources:

      https://rarediseases.info.nih.gov/diseases (to find a rare genetic disease)

      https://www.malacards.org/ (to find genes/pathways associated with disease, use this list as input)

    - *Output*: Display a concise summary of the gene, including:
        - Full name (e.g., "Breast Cancer 1").
        - Biological function (e.g., DNA repair).
        - Associated diseases (e.g., breast cancer).
    - *Source*: Use a mock dataset (e.g., a JSON file with 10-20 genes) or a public API. Here is a list of APIs that will be useful for the task:
        - *Gene Information:*
            - Entrez Utilities:

- https://www.ncbi.nlm.nih.gov/books/NBK25501/
- Example usage for gene information:

```
gene = "EGFR"

BASE_URL = "https://eutils.ncbi.nlm.nih.gov/entrez/eutils/"

esearch_url =
f"{BASE_URL}esearch.fcgi?db=gene&term={gene}[gene]+AND+Homo+sapi
ens[orgn]&retmode=json"

e.g:
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esearch.fcgi?db=ge
ne&term=EGFR[gene]+AND+Homo+sapiens[orgn]&retmode=json

response = requests.get(esearch_url).json()

gene_id = response["esearchresult"]["idlist"][0]

esummary_url =
f"{BASE_URL}esummary.fcgi?db=gene&id={gene_id}&retmode=json"


summary_response = requests.get(esummary_url).json()

gene_info = summary_response["result"][gene_id]

e.g:
https://eutils.ncbi.nlm.nih.gov/entrez/eutils/esummary.fcgi?db=g
ene&id=1956&retmode=json
```

- Similarly you can use OMIM or GeneRIF to get disease associations and functions for genes using the same API as above, just with a different **db** parameter
- Gene Pathway information API's:
    - https://www.kegg.jp/kegg/rest/keggapi.html
    - https://www.wikipathways.org/#download

2. **Gene-Gene Interaction Network**
    - *Visualization*: Generate an interactive network graph showing:
        - The gene of interest as the central node.
        - Connected nodes representing interacting genes (e.g., proteins it binds to or co-expressed genes).
        - Edges indicating interaction type (e.g., physical, regulatory), if feasible.
    - *Similar Genes*: Highlight 3-5 genes that are similar (e.g., members of the same pathway, similarity of effect in other genes' expression).
    - *Implementation*: Use a graph library (e.g., Cytoscape.js, D3.js, or NetworkX) with clickable nodes for details.

- *Source*: Preload a small interaction dataset (e.g., 5-10 genes with mock connections) or simulate based on rules (e.g., "if in same pathway, connect").
    - Gene-gene interactions can be fetched using the following approach:
        - Fetch the pathways a specific gene is involved using https://rest.kegg.jp/link/pathway/{gene_id} (e.g., https://rest.kegg.jp/link/pathway/hsa:7157)
        - For each pathway, fetch the details using https://rest.kegg.jp/get/{pathway_id}/kgml (e.g., https://rest.kegg.jp/get/path:hsa01522/kgml)
        - Fetch the relevant relations between genes (relations that contain the gene_id of interest with the target types: "activation" or "inhibition")
        - Map the fetched gene_ids to the canonical gene name

3. **Drug Repurposing Suggestions**
    - *Output*: For the gene of interest and each similar/interacting gene, list:
        - Known associated drugs (e.g., "Tamoxifen" for BRCA1).
        - Original indication (e.g., "breast cancer").
        - Add a "Repurposing Score" for each drug suggestion, calculated with a simple heuristic (e.g., +10 if drug targets a similar gene, +5 if in the same pathway, -5 if unrelated disease).
    - *Display*: Present in a table or card format next to the network, with clickable drug names for brief info (e.g., mechanism: "inhibitor").
    - *Source*: Use a mock drug-gene database (e.g., CSV with gene-drug pairs) or adapt a small subset from public data (e.g., DrugBank, OpenTargets, ChemBL, etc)

4. **User Interface**
    - *Layout*:
        - Top: Input field for gene name and "Explore" button.
        - Left: Gene overview text.
        - Center: Interactive gene-gene network visualization.
        - Right: Drug repurposing table/cards.
    - *Interactivity*: Clicking a gene node in the network updates the overview and drug list dynamically.
    - *Design*: Clean, intuitive, and responsive (mobile or desktop, your choice).

Technical Requirements

- *Platform*: Mobile or Desktop, based on team skills.
- *Language*: Any (Python, JavaScript, Java, etc.), based on team skills.
- *Libraries*:
    - Graph visualization (e.g., D3.js, PyGraphviz).
    - UI framework (e.g., Bootstrap, Qt).
    - Optional: Basic API integration if using real data.
- *Data*: Preload mock datasets (gene info, interactions, drugs) to keep it hackathon-friendly; no live database required.
- *Performance*: Load and display results within 5 seconds for a single gene query.

Deliverables

1. Working prototype (mobile app, desktop executable, or web app).
2. Brief demo script (2-3 minutes) explaining:
    - How to use the app.
    - A sample gene exploration (e.g., "TP53").
3. Optional: Code repository (e.g., GitHub) with a README.

Success Criteria

- Functional gene input and overview display.
- Visual gene-gene network with at least 5 nodes (including the input gene).
- At least 3 drug suggestions tied to the network, with repurposing ideas.
- Smooth, crash-free demo during judging.

Inspiration from QIAGEN Digital Insights

This app mirrors QDI's mission to turn complex genomic data into actionable insights. Think of it as a mini version of their tools—like the Knowledge Base or CLC Genomics Workbench—applied to drug discovery. Get creative, have fun, and show us how you'd accelerate science!

Explanation: What is Drug Repurposing and Its Value?

Drug repurposing (also called drug repositioning) is the process of identifying new therapeutic uses for existing drugs—drugs that were originally developed and approved for one condition but can be applied to treat a different disease or target. Unlike traditional drug development, which can take over a decade and billions of dollars to bring a new drug to market, repurposing leverages drugs that have already passed safety and pharmacokinetic testing. This significantly reduces time, cost, and risk.

Value of Drug Repurposing:

- *Speed*: Repurposed drugs can reach patients faster since they've already cleared early clinical trial phases.
- *Cost-Effectiveness*: It avoids the high expense of de novo drug discovery.
- *New Insights*: By linking drugs to new gene targets or diseases (e.g., via gene-gene interactions), it uncovers unexpected therapeutic opportunities—like using a cancer drug for a neurodegenerative disease.