Have specific genes been linked/associated with certain diseases?

There are a couple of ways to approach this: either to start with the gene and investigate what diseases it is associated with or start the disease and see if that gene has been associated with it.
There are online databases of genome-wide association studies (GWAS) data which will allow both approaches to the question.

GWAS studies are performed investigating the association between SNPs and a trait, rather than the association of a gene and a trait. The SNPs strongly associated with the trait are then mapped to genes using predicted gene expression from a reference panel. The online databases are searchable by gene. It is worth remembering that this isn't a fool proof system, just because a SNP is close to a gene or even in a gene, it doesn't necessarily mean that the SNP is regulating expression of this gene, FTO is a classic example of this[1].

For more in depth genetic analysis, the SNPs associated with a gene (through either expression of that gene or location in the genome) can be identified through the databases and taken into further downstream analysis (eg. Mendelian Randomization to investigate causality).

With all of these analysis, I would highly recommend discuss your ideas, methods, results with someone within the IEU. These tools are pointers for an initial look up, afer which much validation and discussion should follow.

Online databases:

Open Target Genetics (https://genetics.opentargets.org)
A platform to explore variant-gene-trait associations from genetic data available from UK Biobank and GWAS Catalog. It has a function to search for a gene, variant or trait to identify studies associated with it.

Genotype-Tissue Expression (GTEx) portal (https://gtexportal.org/home/)
The Genotype-Tissue Expression (GTEx) project is an ongoing effort to build a comprehensive public resource to study tissue-specific gene expression and regulation. Samples were collected from 54 non-diseased tissue sites across nearly 1000 individuals, primarily for molecular assays including whole genome sequencing, whole exome sequencing and RNA-Seq. Remaining samples are available from the GTEx Biobank. The GTEx Portal provides open access to data including gene expression, QTLs, and histology images. It relates gene expression from tissues to SNPs.

MR-Base (http://www.mrbase.org):
A platform that integrates a curated database of complete GWAS results (including all genes tested with no restrictions according to statistical significance) with an application programming interface, web app and R packages that automate 2-sample Mendelian randomization. The software includes several sensitivity analyses for assessing the impact of horizontal pleiotropy and other violations of assumptions. The database currently comprises 11 billion single nucleotide polymorphism-trait associations from 1673 GWAS and is updated on a regular basis. It allows for the analysis of an exposure on an outcome.

Ensembl (http://www.ensembl.org/index.html)
Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes,

computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

Expression altas (https://www.ebi.ac.uk/gxa/home)
Gene expression across species and biological conditions. Searchable by gene or by condition.

GWAS central (https://www.gwascentral.org)
GWAS Central (previously the Human Genome Variation database of Genotype-to-Phenotype information) is a database of summary level findings from genetic association studies, both large and small.

Phenome-genotype integrator: PheGenI (https://www.ncbi.nlm.nih.gov/gap/phegeni)
The Phenotype-Genotype Integrator (PheGenI), merges NHGRI genome-wide association study (GWAS) catalog data with several databases housed at the National Center for Biotechnology Information (NCBI), including Gene, dbGaP, OMIM, eQTL and dbSNP. This phenotype-oriented resource, intended for clinicians and epidemiologists interested in following up results from GWAS, can facilitate prioritization of variants to follow up, study design considerations, and generation of biological hypotheses. Users can search based on chromosomal location, gene, SNP, or phenotype and view and download results including annotated tables of SNPs, genes and association results, a dynamic genomic sequence viewer, and gene expression data.

Human Genome Epidemiology (HuGE) Navigator (https://phgkb.cdc.gov/PHGKB/hNHome.action)
HuGE Navigator provides access to a continuously updated knowledge base in human genome epidemiology, including information on population prevalence of genetic variants, gene-disease associations, gene-gene and gene-environment interactions, and evaluation of genetic tests.

eQTLGen Consortium (https://www.eqtlgen.org/index.html)
The eQTLGen Consortium has been set up to identify the downstream consequences of trait-related genetic variants. The consortium incorporates 37 datasets, with a total of 31,684 individuals. You can find the cis-eQTL, trans-eQTL, eQTS and replication results from our forthcoming paper on this website.

Genetic epidemiology methods:

A few genetic epidemiology methodology which may be interesting to laboratory sciences are:

- *Protein GWAS:* GWAS of the plasma proteome[2], metabolites (see review[3]), cytokines[4] are being performed. This will identify SNPs that are associated with the measured and analysed proteins. Pathway analysis will have been performed as follow up analysis to the GWAS.

- *Cell type GWAS:* This is normally the circulating white blood cell count/percentage[5]. More detailed immune cells have been investigated in smaller cohorts[6].

- *Transcriptome-wide association studies (TWAS):* predicts gene expression using gene expression datasets and integrates GWAS data to identify gene-trait associations (see review[7]).

- *Colocalization:* investigates whether potential causal genetic variants (SNPs) are shared between traits or between a SNP, gene expression and a trait.

- *Mendelian Randomization methodologies:* method for assessing causality between an exposure and an outcome

    - *Tissue specific protein MR:* investigates the association between a gene expression and a trait in specific tissues[8].
    - *Plasma proteome MR:* estimates the effect of plasma proteins on phenotypes[9] (http://www.epigraphdb.org/pqtl/).
    - MR resource: *EpiGraphDB:* an analytical platform and database to support data mining in epidemiology. The platform incorporates a graph of causal estimates generated by systematically applying Mendelian randomization to a wide array of phenotypes, and augments this with a wealth of additional data from other bioinformatic sources (http://www.epigraphdb.org).

1. Speakman JR. The "Fat Mass and Obesity Related" (FTO) gene: Mechanisms of Impact on Obesity and Energy Balance.
2. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558(7708):73-79.
3. Kastenmüller G, Raffler J, Gieger C, Suhre K. Genetics of human metabolism: an update. *Hum Mol Genet*. 2015;24(R1):R93-R101.
4. Ahola-Olli A V., Würtz P, Havulinna AS, et al. Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors. *Am J Hum Genet*. 2017;100(1):40-50.
5. Astle WJ, Elding H, Jiang T, Ouwehand WH. The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*. 2016;167:1415-1429.
6. Patin E, Hasan M, Bergstedt J, et al. Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. *Nat Immunol*. 2018;19(3):302-314.
7. Wainberg M, Sinnott-Armstrong N, Mancuso N, et al. Opportunities and challenges for transcriptome-wide association studies. *Nat Genet*. 2019;51(4):592-599.
8. Richardson TG, Hemani G, Gaunt TR, Relton CL, Smith GD. A transcriptome-wide Mendelian randomization study to uncover tissue-dependent regulatory mechanisms across the human phenome. *bioRxiv*. 2019;44(0):563379.
9. Zheng J, Haberland V, Baird D, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *bioRxiv*. 2019:627398.