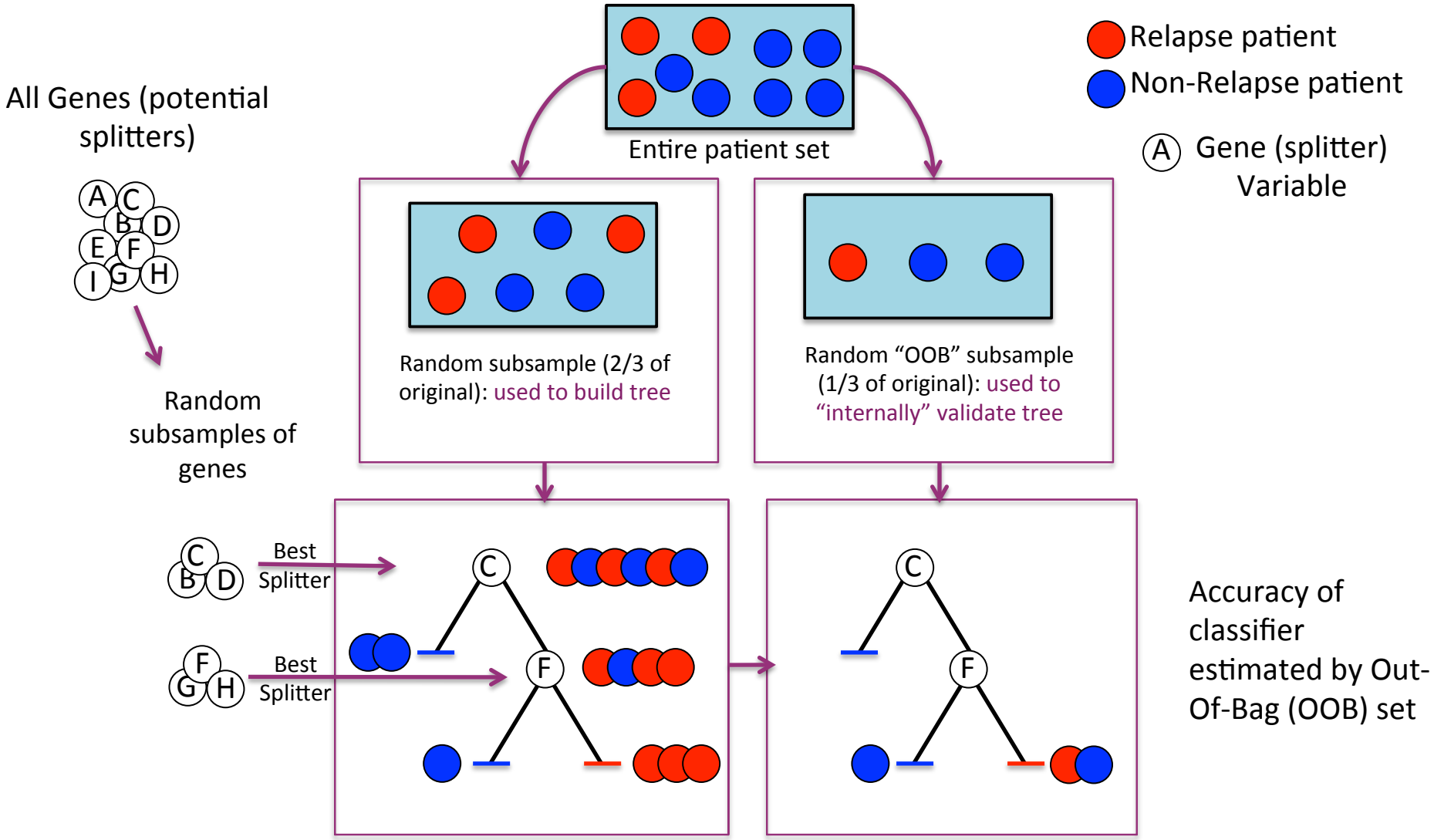
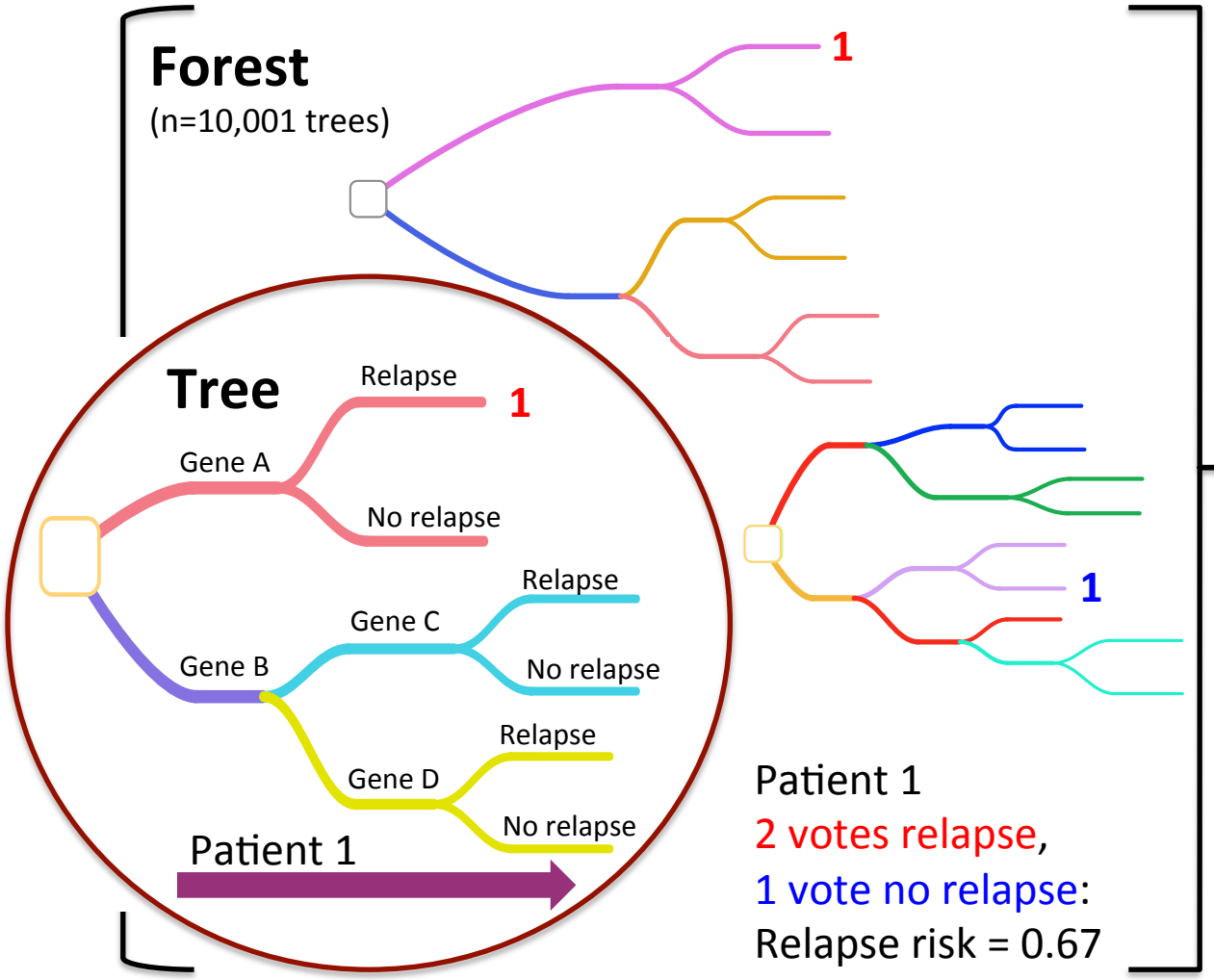
# The Random Forest is built one tree at a time…

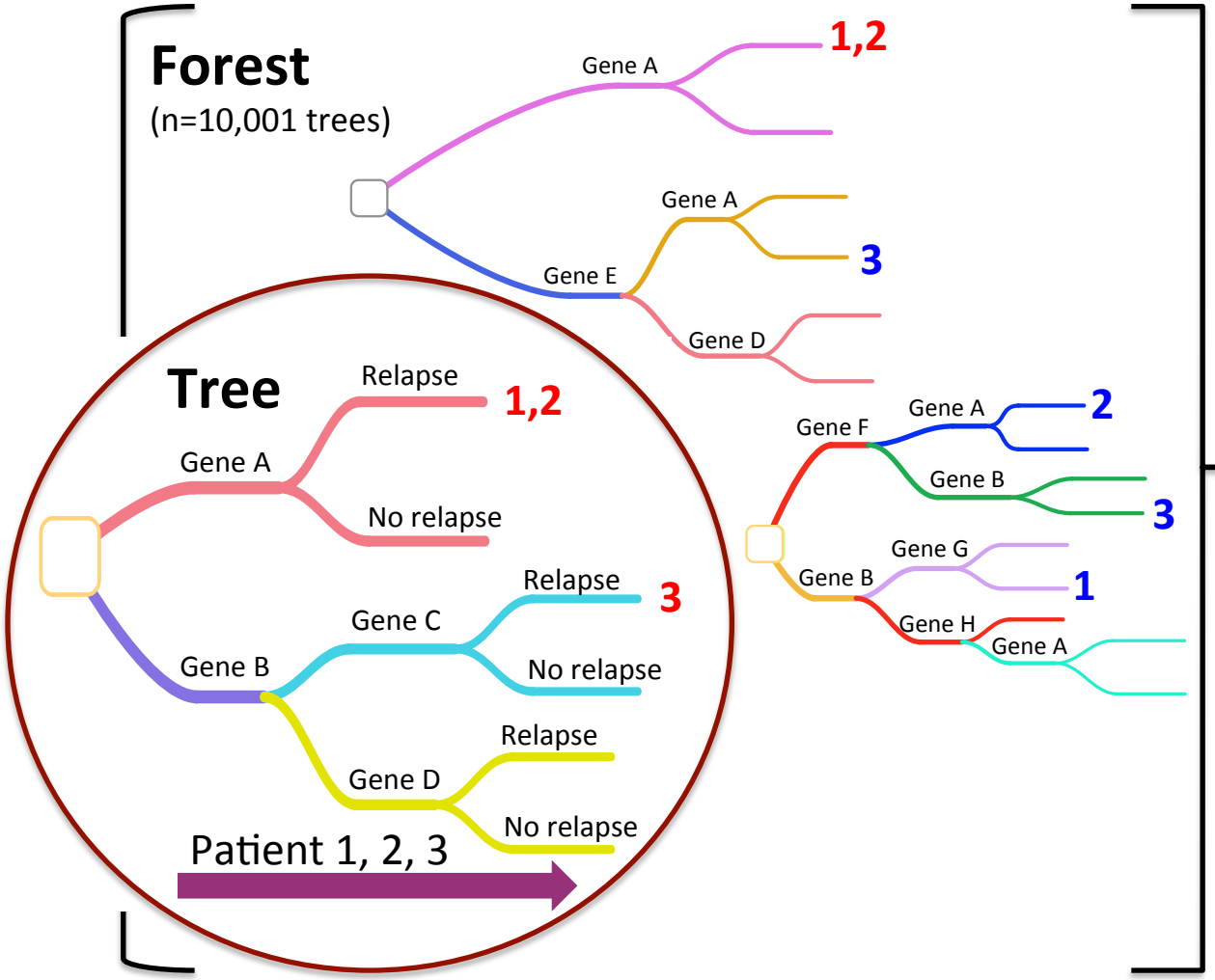


# To predict new patients, each tree gets a vote…




- **Creates a forest of many binary decision trees**
- Each Patient traverses each tree until it reaches a terminal node
- At the terminal node each tree casts a vote (eg. “relapse”); the proportion of relapse votes from all votes is that patient’s predicted relapse risk

# Variable importance is a feature of random forests




- The more often a gene is chosen as a splitter variable, the higher its "Variable Importance" – This can be used to prioritize which genes to select for an assay with limited gene measurements

| Gene | Var. Imp. |
|---|---|
| Gene A | 0.67 |
| Gene B | 0.20 |
| Gene D | 0.13 |
| … | … |