

Supplemental Tables

(in the single excel file attached)

- Table S1. The disease mapping for the current FDA-approved orphan drugs
- Table S2. The disease mapping for the current FDA-designated orphan drugs
- Table S3. The disease category mapped for the current FDA-approved orphan drugs ranked by drug number
- Table S4. The individual diseases mapped for the current FDA-approved orphan drugs ranked by drug number
- Table S5. The disease category mapped for the current FDA-designated orphan drugs ranked by drug number
- Table S6. The individual diseases mapped for the current FDA-designated orphan drugs ranked by drug number
- Table S7. The complete text-mining results of the OMIM for diseases caused by a single gain-of-function mutation
- Table S8. The complete text-mining results of the OMIM for diseases caused by a single loss-of-function mutation and associated with a potential allosteric activator

Category	Gain-of-function	Late onset
Terms	gain-of-function gain of function over-activation over activation overexpression over-expression over expression	adult onset adult on-set adult on set late onset late on-set late on set

Table S9. Terms used to determine gain-of-function and late disease onset in the OMIM disease description.

Category	Structure protein
Gene names	Actin, Actinin, Aggrecan, Cadherin, Collagen, Decorin, Elastin, Fibrinogen, Fibronectin, Heparan, Keratin, Laminin, Mucin, Myosin, Tropomysin, Troponin, Vimentin

Table S10. Terms used to manually determine if a gain-of-function disease is caused by a mutated structure protein