University of Tehran

Identification of Osteogenic Peptides Using Machine Learning Approach

Co-Authors: Taheri-Ledari, Mohammad (University of Tehran), Nowzari-Dalini, Abbas (University of Tehran), Zeynali, Bahman (University of Tehran)

Presenter and first author: Ghobakhloo, Maryam (University of Tehran)

Background

- Considering the critical roles of bone tissue in health (Fig. 1), the treatment of skeletal disorders is therefore of paramount importance. With a fast-growing aging population, osteoporosis becomes a serious health concern. Most treatments for osteoporosis reduce bone resorption, but cannot restore the lost bone structure. Hence, searching for bone-specific osteogenic anabolic drugs is still necessary[1,2].
- In the other hand, bone fractures occurring are associated with impaired healing. Therefore, research surrounding bone tissue engineering and Bioactive Peptides (BAPs) has expanded significantly. Several peptides have shown bone health-promoting effects in vitro and in vivo[3].
- BAPs are more effective and cheaper than proteins and have less toxicity and side effects in the human body. Bioinformatics tools and databases evolved different methods for identifying BAPs[4].
- Up until now, there is not any database that includes Osteogenic Peptides (OPs). Bringing all of these peptides together in a database besides facilitating their accessibility would make a structural and functional study of these peptides convenient. It also helps to the identification of new OPs using bioinformatics methods and tools.

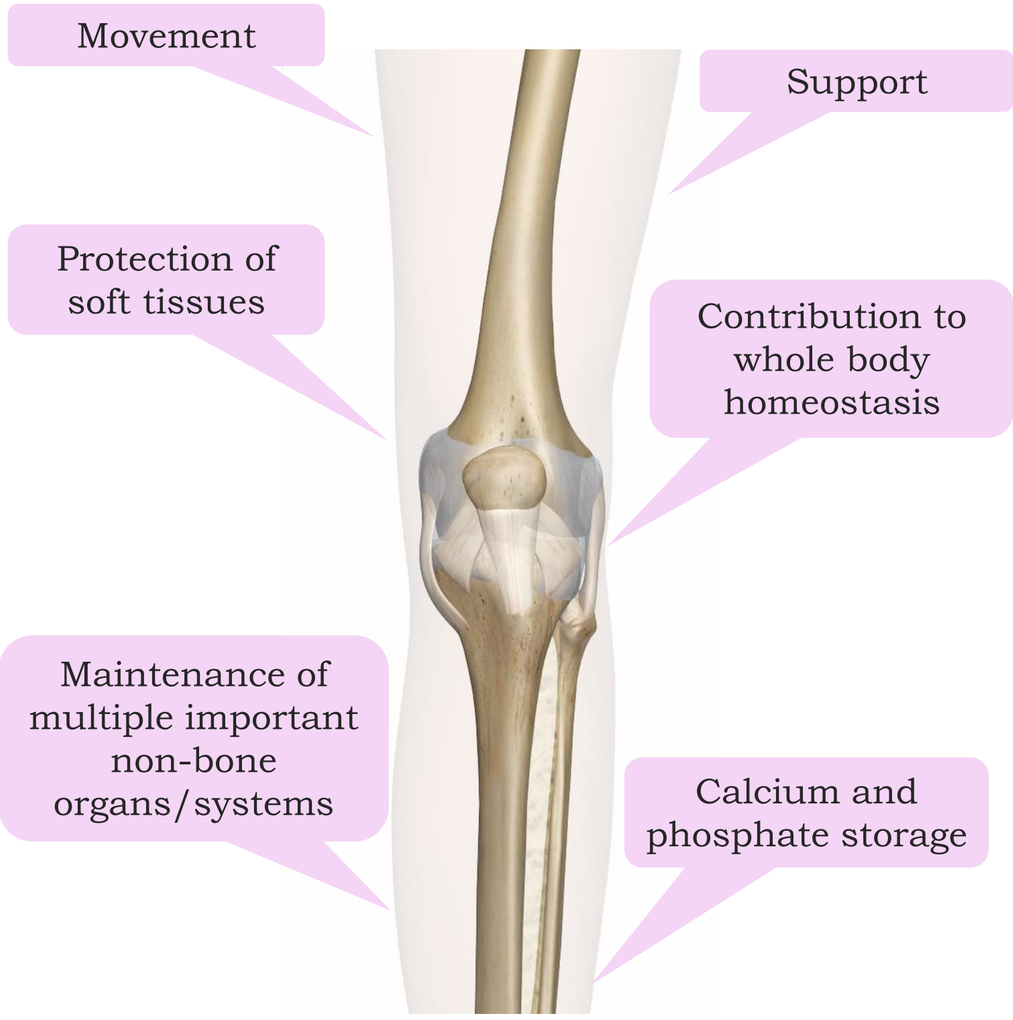

Fig. 1 | The roles of bone tissue in health

Aim

- Developing a database of OPs
- Designing a machine learning-based method to predict OPs within a given protein.
- Discovering peptides with highest osteogenic potentials within caseins

Method

- OPs were collected manually from sets of full-text articles.
- Three pHMMs were designed and trained on these data (Fig. 2).
- The Leave-One-Out-Cross-Validation method was employed for validation.

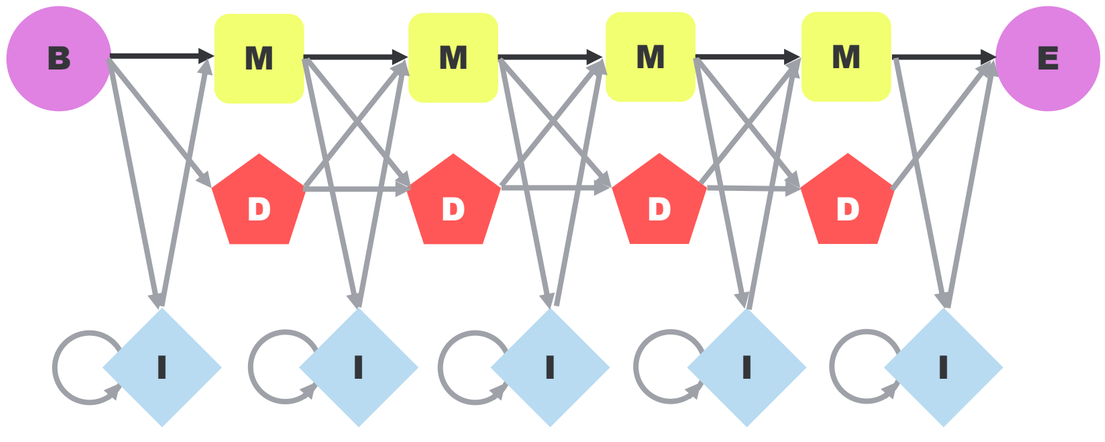

Fig. 2 | A schematic profile Hidden Markov Model
B: Begin, E: End, M: Match, D: Deletion, I: Insertion

Results

- **Database:**

  82 peptides were collected from hundreds of reputable publications.

- **Designing pHMMs:**

  At first, sequences were aligned using the Clustal algorithm. Then, the minimum number of adjacent columns that contained all the sequences of less than or equal to 15 amino acids in length were selected. They were aligned, and a pHMM was designed on them. The same happened for columns with 16 to 20 amino acids in length. The third pHMM was made of remained sequences.

- **Evaluation:**

  Each time, one of the sequences was left out and three pHMMs were designed as described above. Then, 1,000 random sequences with 50 amino acids in length were made and the desired peptide planted between them. Then, all 1000 sequences were given to pHMMs and the precision and recall were calculated. The final F-score was obtained by averaging the precision and recall of all OPs (Table 1).

- **Peptide selection:**

  Finally, three casein-derived peptides were selected using this approach to evaluate their osteogenic potential in the lab.

Table 1 | Results of final cross validation

Recall	Precision	F-score
0.622221	0.521677	0.562682

Conclusions

- Due to the small set of positive data, the final results were not accurate enough. We are still collecting more data to increase accuracy.
- Selected candidate peptides have been synthesized and currently are being tested for their osteogenic inducing ability on human mesenchymal stem cells (Fig. 3).
- Given the importance of bone in health, it seems necessary for researchers in this field to work together to build a coherent database of OPs and develop new methods for identifying them.

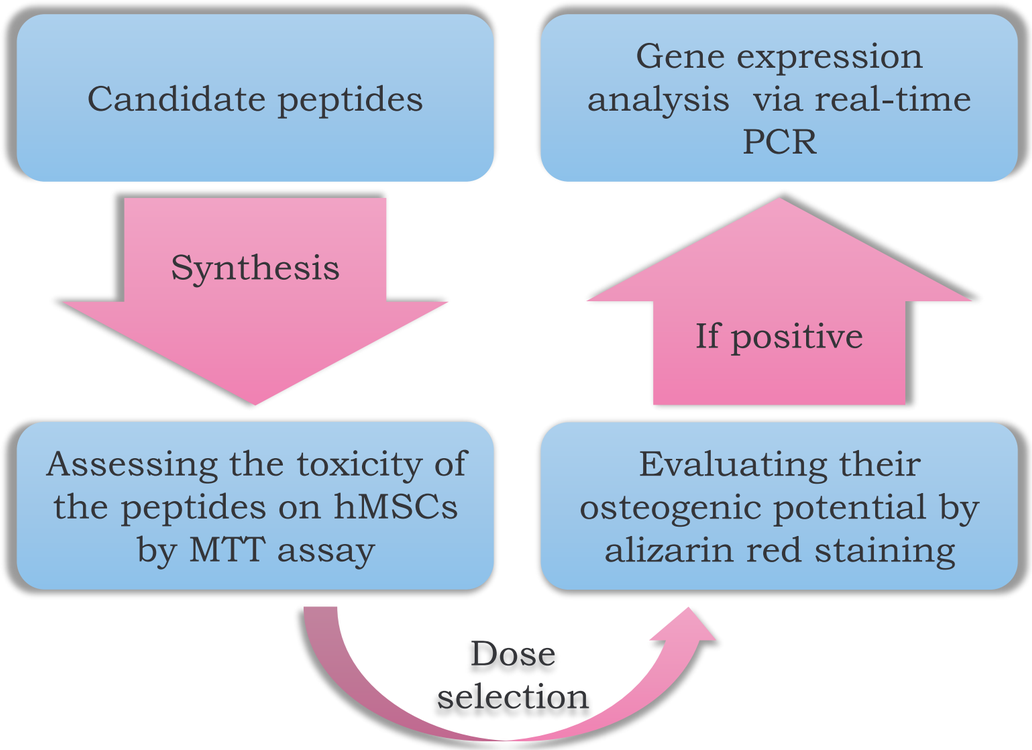

Fig. 3 | Summary of ongoing experiments

References

1. 1. Jiang, M., *et al.* Identification of osteogenic progenitor cell-targeted peptides that augment bone formation. *Nat Commun.* (2020). Doi: 10.1038/s41467-020-17417-9.
2. 2. Amso Z., *et al.* Short Anabolic Peptides for Bone Growth. *Med Res Rev.* (2016). Doi: 10.1002/med.21388.
3. 3. Pountos I., *et al.* The role of peptides in bone healing and regeneration: a systematic review. *BMC Med.* (2016). Doi: 10.1186/s12916-016-0646-y.
4. 4. Tu, M., *et al.* Advancement and prospects of bioinformatics analysis for studying bioactive peptides from food-derived protein: Sequence, structure, and functions. *TrAC Trends in Analytical Chemistry. (2018). Doi: 10.1016/j.trac.2018.04.005*

Maryam.ghobakhloo@ut.ac.ir

University of Tehran
16th Azar St., Enghelab Sq.
1417935840, Tehran, Iran
+98 (21) 61112664
ut.ac.ir