

УДК 575.1

## ГЕНЫ *SEP* КАРТОФЕЛЯ: ИДЕНТИФИКАЦИЯ И АНАЛИЗ ЭКСПРЕССИИ<sup>1</sup>

© 2022 г. Е. А. Рутковская<sup>1</sup>, М. С. Ганчева<sup>1</sup>, \*, М. А. Лебедева<sup>1</sup>, Л. А. Лутова<sup>1</sup>

<sup>1</sup>Санкт-Петербургский государственный университет, Санкт-Петербург, 199034 Россия

\*e-mail: ganchovai@gmail.com

Поступила в редакцию 25.11.2021 г.

После доработки 29.12.2021 г.

Принята к публикации 10.01.2022 г.

*SEP* (C-Terminally Encoded Peptides) – посттрансляционно-модифицируемые пептидные гормоны, играющие важную роль в развитии растений. Пептиды *SEP* известны прежде всего как регуляторы развития корневой системы, активирующиеся в ответ на действие абиотических стрессовых факторов, в частности, на недостаток азота. У картофеля представители семейства *SEP* к настоящему времени не были охарактеризованы. В нашей работе мы идентифицировали 17 генов *SEP* в геноме картофеля (*Solanum tuberosum* L.). С помощью общедоступных данных транскриптомных исследований мы оценили уровни экспрессии генов *SEP* в различных органах картофеля, а также показали, что экспрессия некоторых генов *SEP* зависит от уровня азота в среде. Дальнейшее изучение генов *SEP* у картофеля позволит расширить наше понимание роли сигнальных пептидов в развитии картофеля.

**Ключевые слова:** картофель, гены *SEP*, пептиды *SEP*, азотный сигналинг, азот.

**DOI:** 10.31857/S0016675822060078

Значительную роль в развитии растений играют регуляторные пептиды: передавая информацию от клетки к клетке, они координируют процессы пролиферации и дифференцировки клеток, а также ответы на внешние стимулы. Некоторые регуляторные пептиды растений способны к дальнему транспорту по проводящим тканям и таким образом, задействованы в коммуникации между разными органами растения [1]. К таким пептидам, в частности, относятся пептиды *SEP* (C-Terminally Encoded Peptides); для некоторых представителей этого семейства показана способность транспортироваться по силеме из корня в побег и таким образом опосредовать системный ответ на недостаток азота в почве [2]. Гены семейства *SEP* найдены как среди цветковых, так и голосеменных растений, а также обнаружены у галловых нематод, вызывающих мелойдогиноз [3, 4]. Продуктами генов *SEP* являются белки-предшественники, содержащие сигнальный домен на N-конце, один или несколько консервативных доменов *SEP* и переменный домен между ними [3–5]. В процессе посттрансляционных модификаций белки гидроксилируются по нескольким остаткам пролина и подвергаются протеолизу до образования зрелого пептида, последовательность которого включает лишь консервативный домен *SEP* дли-

ной 15 аминокислот. Пептиды *SEP* регулируют рост корней и побегов растений, а также оказывают влияние на развитие боковых корней и симбиотических клубеньков у бобовых растений [5–7]. Также было показано влияние пептидов *SEP* на процессы поглощения азота растением за счет системной регуляции экспрессии генов, кодирующих нитратные транспортеры [2].

Данные транскриптомных исследований и анализ регуляторных областей генов *SEP* у ряда растений позволили предположить, что роль генов *SEP* в развитии растений не ограничивается их участием в ответе на дефицит азота, и пептиды *SEP*, вероятно, являются важными регуляторами ответа на различные абиотические стрессовые факторы [8]. Однако роль генов *SEP* была изучена лишь у отдельных представителей цветковых растений. В частности, у картофеля к настоящему моменту не были охарактеризованы представители этого семейства генов. В связи с этим, целью нашего исследования стала идентификация и анализ экспрессии генов *SEP* у вида *Solanum tuberosum* L. (*StSEP*). Картофель является важной сельскохозяйственной культурой благодаря его способности образовывать клубни. Известно, что наличие азота в среде является одним из ключевых факторов, регулирующих развитие клубней у картофеля, в частности, высокая концентрация азота в среде задерживает клубнеобразование [9]. Поскольку пептиды *SEP* известны как важные участники ответа на недостаток азота в среде,

<sup>1</sup> Дополнительная информация для этой статьи доступна по doi 10.31857/S0016675822060078 для авторизованных пользователей.

представляется целесообразным изучение возможной роли этих регуляторов в развитии клубней у картофеля.

Для идентификации генов *СЕР* у картофеля мы провели поиск их последовательностей в геноме картофеля SolTub 3.0 (Ensembl:GCA\_000226075.1) с использованием всех известных генов *СЕР*, выявленных ранее в работе [3], с помощью алгоритма tblastx. Также был произведен поиск доменов *СЕР* в последовательностях картофеля с помощью программ Multiple Expectation Maximization for Motif Elicitation (MEME) (v. 5.4.1), Find Individual Motif Occurrences (FIMO) (v. 5.4.1) и Motif Alignment & Search Tool (MAST) (v. 5.4.1) [10] с использованием доменов *СЕР*, обнаруженных в исследовании [3]. Полученные последовательности были проанализированы на наличие сигнального домена на N-конце с помощью программы SignalP (v. 5.0) [11]. Результатом проделанной работы стало обнаружение 17 последовательностей St*СЕР* (Прилож., табл. 1). Локализация генов *StСЕР* на хромосомах была визуализирована с помощью программы MapInspect (<http://www.softsea.com/download/MapInspect.html>). При этом оказалось, что девять генов *StСЕР* располагаются на одной хромосоме (рис. 1,а). Среди них: гены *StСЕР3* и *StСЕР9* располагаются в непосредственной близости друг от друга, а гены *StСЕР1*, 6, 7, 13 образуют большой кластер из tandemно расположенных генов.

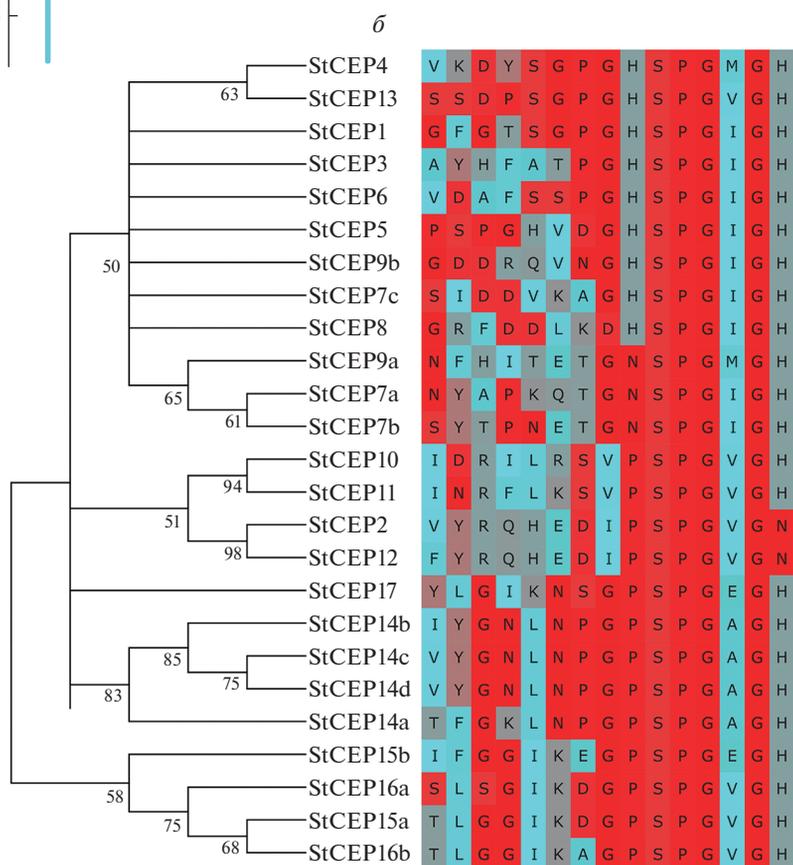
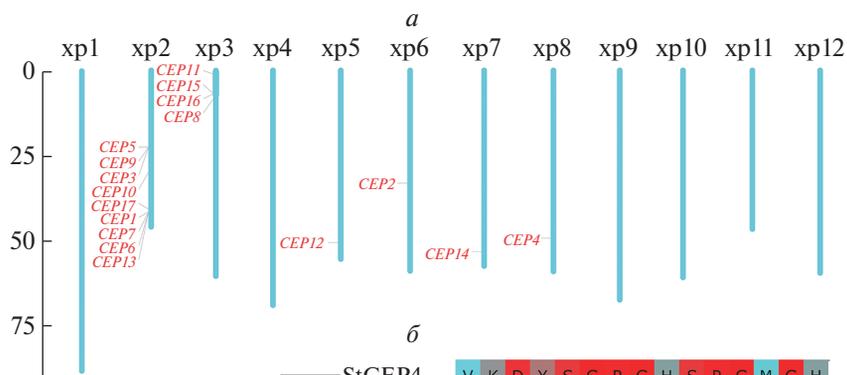
На основании нуклеотидных последовательностей были определены аминокислотные последовательности белков St*СЕР*. Аминокислотные последовательности *СЕР* выравнивали в программе Ugene [12], и на основании выравнивания было построено филогенетическое дерево методом максимального правдоподобия в программе MEGA11 с bootstrap-поддержкой в 1000 реплик. Среди белков St*СЕР* оказались как белки, содержащие только один домен *СЕР* (St*СЕР1*–13), так и мультидоменные белки, содержащие два (в случае St*СЕР15*, 16) или четыре (в случае St*СЕР14*) консервативных домена *СЕР*. Так оказалось, что белок St*СЕР14* содержит четыре практически идентичных по последовательности домена *СЕР* (рис. 1,б). Последовательности доменов *СЕР* в пределах мультидоменных белков St*СЕР15* и St*СЕР16* характеризуются более выраженными различиями по аминокислотной последовательности, при этом отдельные домены из разных мультидоменных белков St*СЕР*, обозначенные нами как St-

*СЕР15a* и St*СЕР16b*, демонстрируют высокое сходство по последовательности и различаются только по одной аминокислоте (рис. 1,б). Ранее в работе Н.А. Ogilvie с соавт. было отмечено, что консенсусная аминокислотная последовательность пептидов *СЕР* у представителей семейства пасленовые отличается от последовательностей *СЕР*, характерных для других семейств растений [3]. В связи с этим можно предположить, что конкретные функции выявленных нами пептидов *СЕР* у картофеля могут отличаться от описанных ранее для других видов растений.

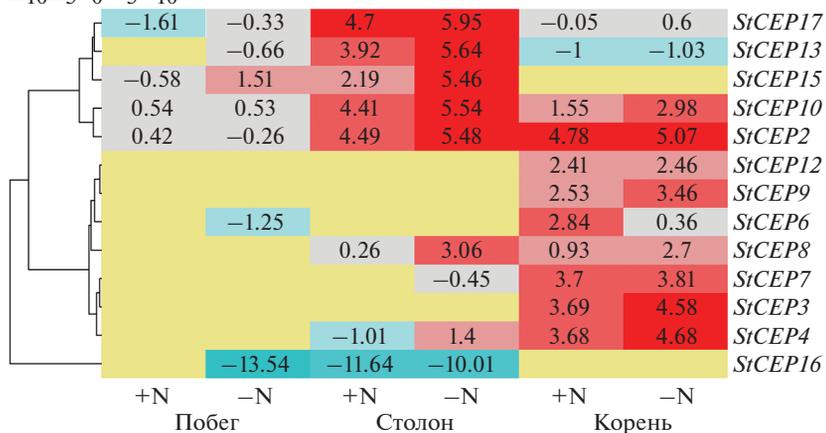
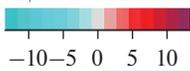
Мы проанализировали транскриптомы листьев и корней трех сортов картофеля, выращенных в почве с азотом (3.3 кг/100 м<sup>2</sup> мочевины) и без него, доступные в базе данных NCBI (проект номер PRJNA511561), данные представлены в трех повторностях. Исследователями в работу были взяты сорта картофеля, различающиеся по эффективности поглощения и использования азота. Эффективность поглощения азота соответствует доле доступного азота в почве, которая была поглощена растением. Эффективность использования рассчитывается как урожай клубней, деленный на накопленный азот во всем растении. Так, картофель сорта “Yanshu 4” эффективно поглощает азот, но имеет низкую эффективность его использования, сорт “Xiabodi” (cv. Sheredy) характеризуется средними показателями эффективности поглощения и использования, а сорт “Chunshu 4” имеет малую эффективность поглощения и высокую – использования [13].

Так как выявленные нами гены *StСЕР* не аннотированы в геномах картофеля, мы заново проводили анализ данных секвенирования. Качество прочтений проверяли с помощью программы FastQC (v. 0.11.9) [14] и визуализировали отчеты программы FastQC с помощью программы MultiQC [15]. Из сырых данных секвенирования удаляли адаптеры с помощью программы Trimmomatic-0.39 [16]. Подсчет количества прочтений на транскриптом картофеля, в который нами были включены транскрипты генов *StСЕР*, проводили с помощью программы kallisto (v. 0.46.2) [17]. Дифференциально экспрессирующиеся гены ( $p\_value < 0.01$ ,  $\log_2\text{FoldChange} > |1|$ , минимальное количество прочтений – десять) находили с помощью пакета DeSeq2 (v. 1.34.0) [18] в программе R 4.1.1 [19] (Прилож., табл. 2). Среди генов *StСЕР* дифференциально экспрессировались всего два гена – *StСЕР15* и *StСЕР17*. У сорта “Yanshu 4”, который эффективно поглощает азот, но имеет низкую эффективность использования азота, значимое снижение количества транскриптов гена *StСЕР17*

**Рис. 1.** Характеристика генов и пептидов *СЕР* картофеля. а – локализация на хромосомах генов *СЕР* картофеля, б – филогенетическое дерево и выравнивание аминокислотных последовательностей пептидов *СЕР* картофеля *Solanum tuberosum* (St*СЕР*). Латинские буквы (а–д) рядом с названиями пептидов обозначают различные домены *СЕР*, выявленные в пределах одного белка-предшественника. в – Log<sub>2</sub> значения TPM (TranscriptsPerMillion) для генов *СЕР* картофеля, растущих в условиях большого (4 мМ, +N) и малого (0.2 мМ, –N) содержания азота в среде. Желтым цветом обозначены TPM равные 0.



Цветовая шкала



наблюдали в корнях в условиях наличия азота ( $\log_2\text{FoldChange} = -1.19$ ,  $\text{padj} < 0.025$ ), и, соответственно, при недостатке азота уровень экспрессии этого гена в корнях был выше.

Известно, что у резуховидки экспрессия ряда генов *CEP* активируется в корнях при недостатке азота, и пептиды *CEP* увеличивают эффективность поглощения азота корнями за счет их участия в активации экспрессии генов нитратных транспортеров. Более высокий уровень экспрессии гена *StCEP17* в корнях сорта “Yanshu 4” в условиях нехватки азота может свидетельствовать о том, что пептид *StCEP17* также может регулировать поглощение нитрата при недостатке азота. При этом у сорта “Chunshu 4”, который характеризуется малой эффективностью поглощения и высокой эффективностью использования азота, выявлено значимое увеличение количества транскриптов гена *StCEP15* ( $\log_2\text{FoldChange} = 2.77$ ,  $\text{padj} < 0.003$ ) в листьях в условиях наличия азота. Тот факт, что экспрессия *StCEP15* возрастает в листьях, но не в корнях (при наличии азота в среде) позволяет предположить, что соответствующий пептид вероятно не задействован в передаче сигнала об изменении уровня азота в среде от корней в побег; и он может выполнять иные функции в ответе на изменение содержания азота в среде.

В работе [20] были также секвенированы транскриптомы картофеля, растущего на среде с высоким (4 мМ) и низким (0.2 мМ) содержанием азота. Азот был представлен в двух формах  $\text{NO}_3^-$  и  $\text{NH}_4^+$ . В анализ были взяты разные органы картофеля: побеги, столоны и корни картофеля сорта “Kufri Jyoti”, который выращивали в условиях аэропоники [20]. В ходе данного исследования были секвенированы транскриптомы для образцов из одной повторности (проект номер PRJNA529319). Мы проанализировали данные, полученные в работе [20], по алгоритму, описанному выше. В результате этого получили значения уровней экспрессии генов, выраженные в виде единиц TPM (Transcripts Per Million) (доп. табл. 3), сгенерированные программой kallisto (v. 0.46.2) [17]. Гены *StCEP1*, *StCEP5*, *StCEP11* и *StCEP14* мы исключили из анализа в связи с отсутствием транскриптов этих генов во всех анализируемых органах. Для остальных генов была построена тепловая карта уровней экспрессии генов, выраженных в единицах  $\log_2$  от значения TPM, что было сделано с помощью функции heatmap.2 [21] из пакета gplots (<https://github.com/talgalili/gplots>) в программе R 4.1.1 [19] (рис. 1, в).

Гены были кластеризованы в соответствии с характером их экспрессии. Интересно отметить, что пять генов *StCEP* (*StCEP2*, *10*, *13*, *15*, *17*) продемонстрировали высокий уровень экспрессии в столоне картофеля, однако большая часть генов *StCEP* все же экспрессировалась в корнях, при этом ген *StCEP2* экспрессировался на относительно высоких уровнях как в столонах, так и в корнях. Было обнаружено изменение экспрессии не-

которых генов *StCEP* при выращивании растений при разных условиях среды – с высоким и низким содержанием азота. Так, в побеге увеличивается количество транскриптов генов *StCEP15* и *StCEP17* и уменьшается экспрессия *StCEP2* при сравнении транскриптомов растений, растущих с малым содержанием азота, и транскриптомов растений, растущих с высоким содержанием азота. При этом результаты двух проанализированных транскриптомов схожи. Так, экспрессия гена *StCEP17* повышается в корнях растений сорта “Yanshu 4” и “Kufri Jyoti” в условиях нехватки азота. В то же время экспрессия гена *StCEP17* значимо не изменялась у сортов “Xiabodi” и “Chunshu 4”. Эти сорта отличаются по эффективности поглощения и использования азота, и пептиды *CEP*, как сигналы о содержании азота в среде, могут быть задействованы в процессах, связанных с определением азотного статуса растений.

Таким образом, наша работа является вкладом в исследование еще малоизученного семейства пептидов *CEP*. Мы идентифицировали 17 генов *StCEP* у картофеля и проанализировали их экспрессию в транскриптомах четырех сортов картофеля, различающихся по эффективности использования поступающего из среды азота, выращенных как в богатых азотом почвах, так и при его недостатке. Нами были выявлены гены *StCEP*, экспрессия которых изменялась в зависимости от наличия азота в среде. Можно предположить, что кодируемые этими генами пептиды выступают в роли сигнальных молекул, задействованных в ответных реакциях на содержания азота в почве. Также мы обнаружили экспрессию ряда генов *StCEP* в столонах картофеля, что позволяет предположить их возможное участие в процессах клубнеобразования.

Работа выполнена при поддержке гранта РФФИ 19-016-00177.

Настоящая статья не содержит каких-либо исследований с использованием в качестве объекта животных.

Настоящая статья не содержит каких-либо исследований с участием в качестве объекта людей.

Авторы заявляют, что у них нет конфликта интересов.

## СПИСОК ЛИТЕРАТУРЫ

1. Lebedeva M.A., Yashenkova Y.S., Dodueva I.E., Lutova L.A. Molecular dialog between root and shoot via regulatory peptides and its role in systemic control of plant development // Russ. J. Plant Physiology. 2020. V. 67. № 6. P. 985–1002. <https://doi.org/10.1134/S1021443720060114>
2. Tabata R., Sumida K., Yoshii T. et al. Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling // Science. 2014. V. 17. № 346. P. 343–346. <https://doi.org/10.1126/science.1257800>
3. Ogilvie H.A., Imin N., Djordjevic M.A. Diversification of the C-Terminally Encoded Peptide (CEP) gene fam-

- ily in angiosperms, and evolution of plant-family specific CEP genes // *BMC Genomics*. 2014. V. 15. № 870. P. 1–15. <https://doi.org/10.1186/1471-2164-15-870>
4. Roberts I., Smith S., De Rybel B. et al. The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development // *J. Exp. Bot.* 2013. V. 64. № 17. P. 5371–5381. <https://doi.org/10.1093/jxb/ert331>
  5. Delay C., Imin N., Djordjevic M.A. Regulation of Arabidopsis root development by small signaling peptides // *Front. Plant Sci.* 2013. V. 4. № 352. P. 1–6. <https://doi.org/10.3389/fpls.2013.00352>
  6. Ohyama K., Ogawa M., Matsubayashi Y. Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis // *Plant J.* 2008. V. 55. № 1. P. 152–160. <https://doi.org/10.1111/j.1365-313X.2008.03464.x>
  7. Imin N., Mohd-Radzman N.A., Ogilvie H.A., Djordjevic M.A. The peptide-encoding *CEP1* gene modulates lateral root and nodule numbers in *Medicago truncatula* // *J. Exp. Bot.* 2013. V. 64. № 17. P. 5395–5409. <https://doi.org/10.1093/jxb/ert369>
  8. Aggarwal S., Kumar A., Jain M. et al. C-terminally encoded peptides (CEPs) are potential mediators of abiotic stress response in plants // *Physiol. Mol. Biol. Plants*. 2020. V. 26. № 10. P. 2019–2033. <https://doi.org/10.1007/s12298-020-00881-4>
  9. Ewing E.E., Struik P.C. Tuber formation in potato: Induction, initiation, and growth // *Hortic. Rev.* 1992. V. 14. P. 89–198. <https://doi.org/10.1002/9780470650523.ch3>
  10. Bailey T.L., Johnson J., Grant C.E., Noble W.S. The MEME Suite // *Nucl. Ac. Res.* 2015. V. 43(W1). P. 39–49. <https://doi.org/10.1093/nar/gkv416>
  11. Almagro Armenteros J.J., Tsirigos K.D., Sønderby C.K. SignalP 5.0 improves signal peptide predictions using deep neural networks // *Nat. Biotechnol.* 2019. V. 37. № 4. P. 420–423. <https://doi.org/10.1038/s41587-019-0036-z>
  12. Okonechnikov K., Golosova O., Fursov M. UGENE team. Unipro UGENE: A unified bioinformatics toolkit // *Bioinformatics*. 2012. V. 28. P. 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
  13. Zhang J., Wang Y., Zhao Y. et al. Transcriptome analysis reveals Nitrogen deficiency induced alterations in leaf and root of three cultivars of potato (*Solanum tuberosum* L.) // *PLoS One*. 2020. V. 15. № 10. P. 1–17. <https://doi.org/10.1371/journal.pone.0240662>
  14. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. 2010. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
  15. Ewels P., Magnusson M., Lundin S., Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report // *Bioinformatics*. 2016. V. 32(19). P. 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
  16. Bolger A.M., Lohse M., Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data // *Bioinformatics*. 2014. V. 30. № 15. P. 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
  17. Bray N.L., Pimentel H., Melsted P., Pachter L. Near-optimal probabilistic RNA-seq quantification // *Nat. Biotechnol.* 2016. V. 34. № 5. P. 525–527. <https://doi.org/10.1038/nbt.3519>
  18. Love M.I., Huber W., Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 // *Genome Biol.* 2014. V. 15. № 550. P. 1–21. <https://doi.org/10.1186/s13059-014-0550-8>
  19. R Core Team. R: A language and environment for statistical computing // R Foundation for Stat. Computing. Vienna, Austria: 2021. <https://www.R-project.org>
  20. Tiwari J.K., Buckseth T., Zinta R. et al. Transcriptome analysis of potato shoots, roots and stolons under nitrogen stress // *Sci. Rep.* 2020. V. 10. № 1152. P. 1–18. <https://doi.org/10.1038/s41598-020-58167-4>
  21. Warnes G.R., Bolker B., Bonebakker L. et al. Gplots: Various R Programming Tools for Plotting Data // R package version 3.1.1. 2020. <https://CRAN.R-project.org/package=gplots>

## Identification and Expression Analysis of CEP Genes in Potato

E. A. Rutkovskaya<sup>a</sup>, M. S. Gancheva<sup>a</sup>, \*, M. A. Lebedeva<sup>a</sup>, and L. A. Lutova<sup>a</sup>

<sup>a</sup>Saint-Petersburg State University, Saint-Petersburg, 199034 Russia

\*e-mail: ganchovai@gmail.com

C-terminally encoded peptide (CEP), post-translationally modified peptide hormones, play important roles in plant development. CEP peptides are known as regulators of root system in response to abiotic stress factors, in particular, to a nitrogen deficiency. However, the members of the CEP family in potato have not yet been characterized. In our work, we identified 17 CEP genes in the potato (*Solanum tuberosum*) genome. Using available transcriptomic data, we estimated the expression levels of the CEP genes in different organs, and found that the expression levels of some of the CEP genes depend on the nitrogen amount in the media. Further study of the CEP genes in potato will expand our understanding of the role of these signal peptides in potato development.

**Keywords:** potato, CEP peptides, peptide hormones, nitrogen.