

(19)



Евразийское
патентное
ведомство

(21) 202091701 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2021.01.29

(51) Int. Cl. C12N 15/82 (2006.01)
A01H 5/00 (2006.01)

(22) Дата подачи заявки
2010.07.20

(54) УВЕЛИЧЕНИЕ БИОМАССЫ ТРАНСГЕННЫХ РАСТЕНИЙ

(31) 61/226,969

(32) 2009.07.20

(33) US

(62) 201791502; 2010.07.20

(71) Заявитель:
СЕРЕС, ИНК. (US)

(72) Изобретатель:

У Чуань-Инь, Ким Хан-Сук,
Магпантай Джерард, Чжоу Фасун,
Соса Джулисса, Надзан Грег, Пеннелл
Роджер И., Эчирилоуи Мирсеа, Ван Уи
(US)

(74) Представитель:

Медведев В.Н. (RU)

(57) В описании изобретения раскрываются методы и материалы для модуляции уровней биомассы у растений. Например, в описании раскрываются нуклеиновые кислоты, которые кодируют модулирующие биомассу полипептиды, а также методы использования таких нуклеиновых кислот для трансформации растительных клеток. Также в описании раскрываются растения, имеющие увеличенный уровень биомассы, и растительные продукты, полученные из растений, имеющих увеличенный уровень биомассы.

A1

202091701

202091701

A1

УВЕЛИЧЕНИЕ БИОМАССЫ ТРАНСГЕННЫХ РАСТЕНИЙ

ПЕРЕКРЕСТНЫЕ ССЫЛКИ НА РОДСТВЕННЫЕ ЗАЯВКИ

В настоящей заявке изложены преимущества предварительной заявки США, серийный номер 61/226,969, поданной 20 июля 2009 г. Раскрытие предварительной заявки включено путем ссылки в полном объеме.

ОБЛАСТЬ ТЕХНИКИ

Настоящий документ касается методов и материалов, применяемых при модуляции уровня биомассы растений. В настоящем документе приведены примеры растений, увеличивших уровень биомассы, а также материалов и методов, с помощью которых было достигнуто увеличение биомассы растений и продуктов растительного происхождения.

ПРЕДПОСЫЛКИ СОЗДАНИЯ ИЗОБРЕТЕНИЯ

Настоящее изобретение относится к методам увеличения биомассы растений и самих растений, которые получены с их помощью. Растения с увеличенной биомассой и/или биомассой повышенной ценности пригодны для использования в сельском хозяйстве, садоводстве, в качестве биомассы для преобразования энергии, производства бумаги, производства продуктов растительного происхождения и других отраслей промышленности. В частности, существует потребность в увеличении биомассы отдельных сельскохозяйственных культур, используемых в качестве источника энергии, например *Panicum virgatum* L. (просо), *Miscanthus x giganteus* (мискантус), *Sorghum* sp. и *Saccharum* sp. (сахарный тростник). На протяжении истории человечества возможность использования растительной биомассы в качестве пищи и топлива имела существенное значение для поддержания и увеличения численности популяции. Ученые непрерывно ищут пути повышения ценности биомассы сельскохозяйственных культур. Большое количество исследований, относящихся к увеличению биомассы растений, в частности отдельных сельскохозяйственных культур, используемых в качестве источника энергии, свидетельствует об огромном значении, которое придается

обеспечению населения надежными источниками энергии. Острая необходимость в создании надежных и стабильных источников растительной биомассы, используемой для получения энергии, диктуется текущими событиями, такими как рост цен на нефть. Объем биомассы, продуцируемой растениями, является количественным признаком, на который оказывает влияние ряд биохимических путей. Существует потребность в молекулярно-генетических методах более быстрого получения растений с увеличенной биомассой. Кроме того, существует необходимость в получении видов растений, которые растут более эффективно и продуцируют большее количество биомассы в различных географических и/или климатических условиях. Такие методы желательно применять к нескольким видам растений (Zhang *et al.*, *Plant Physiol.* 135: 615-621 (2004)). Несмотря на определенный прогресс в молекулярно-генетических методах, также существует необходимость в определении специфических генов и/или последовательностей, которые можно использовать для эффективного увеличения биомассы растений.

КРАТКОЕ ИЗЛОЖЕНИЕ

В настоящем документе приведены методы и материалы, относящиеся к растениям с модулированным уровнем биомассы. Например, в настоящем документе представлены трансгенные растения и растительные клетки с увеличенным уровнем биомассы, нуклеиновые кислоты, используемые для получения трансгенных растений и растительных клеток с увеличенным уровнем биомассы, методы увеличения биомассы растений и методы производства растительных клеток, которые можно использовать для получения растений с увеличенным уровнем биомассы. Такие растения и растительные клетки можно выращивать, например, для получения растений с увеличенной высотой, увеличенным количеством побегов или увеличенным сухим весом. Растения с увеличенной биомассой могут быть полезны для производства биомассы, которая может использоваться как в качестве пищи для человека, так и в качестве корма для животных. Растения с увеличенной биомассой могут быть полезны при превращении такой биомассы в жидкое топливо (например, этанол) или другие химические вещества либо

могут использоваться в качестве термохимического топлива.

В настоящем документе представлены методы получения растений с увеличенной биомассой. С одной стороны, метод включает в себя выращивание растительной клетки с экзогенной нуклеиновой кислотой. Экзогенная нуклеиновая кислота содержит регуляторную область, которая функционально связана с нуклеотидной последовательностью, кодирующей полипептид. Показатель аминокислотной последовательности полипептида по скрытой марковской модели (Hidden Markov Model, НММ) больше 130, 340, 530, 120, 635, 65, 100, 480, 145, 280 или 1000, если использовать показатели НММ, полученные из аминокислотных последовательностей, показанных, соответственно, на фиг. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или 11. Уровень биомассы растения отличается от соответствующего уровня биомассы контрольного растения, которое не содержит экзогенной нуклеиновой кислоты.

С другой стороны, метод включает в себя выращивание растительной клетки, содержащей экзогенную нуклеиновую кислоту. Экзогенная нуклеиновая кислота содержит регуляторную область, которая функционально связана с нуклеотидной последовательностью, кодирующей полипептид, имеющий идентичность последовательности 80% или более по отношению к аминокислотной последовательности, представленной в последовательностях SEQ ID NO:1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302,

304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697. Растение, полученное из растительной клетки, может быть использовано для генерации растения, уровень биомассы которого отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота.

С другой стороны, метод включает в себя выращивание растительной клетки, содержащей экзогенную нуклеиновую кислоту. Экзогенная нуклеиновая кислота содержит регуляторную область, которая функционально связана с нуклеотидной последовательностью, имеющей идентичность последовательности 80% или более по отношению к нуклеотидной последовательности (или ее фрагменту), представленной в последовательностях SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233,

235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696. Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота.

В настоящем документе представлены методы модуляции уровня биомассы растения. С одной стороны, метод заключается во введении в растительную клетку экзогенной нуклеиновой кислоты, имеющей регуляторную область, которая функционально связана с нуклеотидной последовательностью, кодирующей полипептид. Показатель аминокислотной последовательности полипептида НММ больше 130, 340, 530, 120, 635, 65, 100, 480, 145, 280 или 1000, если использовать показатели НММ, полученные из аминокислотных последовательностей, показанных, соответственно, на фиг. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или 11. Уровень биомассы растение, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота.

В некоторых вариантах показатель НММ аминокислотной последовательности полипептида превышает 340, если использовать показатели НММ, полученные из аминокислотных последовательностей, показанных на фиг. 2, в которых полипептид содержит ДНК-связывающий домен «цинковые пальцы», имеющий идентичность последовательности не менее 60% (например, 65, 70,

75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 130-192 последовательности SEQ ID NO: 263, или ДНК-связывающий домен «цинковые пальцы», обозначенный в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 530, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 3, в которых полипептид содержит домен фитохелатин-синтетаза, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 44-208 последовательности SEQ ID NO: 117, или домены фитохелатин-синтетаза, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 120, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 4, в которых полипептид содержит домен AP2, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 32-83 последовательности SEQ ID NO: 1, или домены AP2, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 635, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 5, в которых полипептид содержит домен аминотрансфераза класса I и II, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 88-453 последовательности SEQ ID NO: 645, или домены аминотрансфераза класса I и II, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 100, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 7, в которых полипептид содержит ДНК-связывающий домен семейства Myb, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 13-62

последовательности SEQ ID NO: 323, или ДНК-связывающие домены семейства Myb, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 480, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 8, в которых полипептид содержит домен альфа/бета-складки гидролазы, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 35-257 последовательности SEQ ID NO: № 595.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 145, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 9, в которых полипептид содержит домен «фактор быстрого подщелачивания» (ФБП), имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 57-129 последовательности SEQ ID NO: 77, или домены ФБП, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 280, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 10, в которых полипептид содержит домен белка (DUF640) с неизвестной функцией, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 19-152 последовательности SEQ ID NO: 209, или домены DUF640, обозначенные в перечне последовательностей.

В некоторых вариантах показатель HMM аминокислотной последовательности полипептида превышает 1000, если использовать показатели HMM, полученные из аминокислотных последовательностей, показанных на фиг. 11, в которых полипептид содержит домен семейства POT, имеющий идентичность последовательности не менее 60% (например, 65, 70, 75, 80, 85, 90, 95, 99 или 100%) по отношению к остаткам 100-509 последовательности SEQ ID NO: 426, или домены семейства POT,

обозначенные в перечне последовательностей.

С другой стороны, метод включает внедрение в растительную клетку экзогенной нуклеиновой кислоты, которая содержит регуляторную область, функционально связанную с нуклеотидной последовательностью, кодирующей полипептид, с идентичностью последовательности 80% или более по отношению к аминокислотной последовательности, представленной в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613,

615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697. Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота. Полипептид в любом из вышеуказанных методов может иметь аминокислотную последовательность, представленную в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560,

562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697.

С другой стороны, метод включает внедрение в растительную клетку экзогенной нуклеиновой кислоты, которая содержит регуляторную область, функционально связанную с нуклеотидной последовательностью, которая на 80% или более идентична нуклеотидной последовательности, представленной в последовательностях SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696, или ее фрагменту. Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота.

В настоящем документе приведены данные о растительных клетках, содержащих экзогенную нуклеиновую кислоту. С одной стороны, экзогенная нуклеиновая кислота содержит регуляторную область, которая функционально связана с нуклеотидной последовательностью, кодирующей полипептид. Показатель НММ аминокислотной последовательности полипептида больше 130, 340, 530, 120, 635, 65, 100, 480, 145, 280 или 1000, если использовать показатели НММ, полученные из аминокислотных последовательностей, показанных на фиг. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или 11. Уровень биомассы растения отличается от соответствующего уровня биомассы контрольного растения, которое не содержит экзогенной нуклеиновой кислоты. С другой стороны, экзогенная нуклеиновая кислота содержит регуляторную область, которая функционально связана с нуклеотидной последовательностью, кодирующей полипептид, с идентичностью последовательности 80% или более по отношению к аминокислотной последовательности, выбранной из группы последовательностей SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409,

411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697.

Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота. С другой стороны, экзогенная нуклеиновая кислота содержит регуляторную область, функционально связанную с нуклеотидной последовательностью с идентичностью последовательности 80% или более по отношению к нуклеотидной последовательности, выбранной из группы последовательностей SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514,

520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696, или ее фрагменту. Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует экзогенная нуклеиновая кислота. Также представлено трансгенное растение, содержащее такую растительную клетку. Кроме того, представлена растительная биомасса или семенная продукция. Данная продукция включает растительную или эмбриональную ткань трансгенного растения, описанного в настоящем документе.

Также представлены изолированные нуклеиновые кислоты. С одной стороны, изолированная нуклеиновая кислота включает в себя нуклеотидную последовательность, имеющую идентичность последовательности 80% или более по отношению к нуклеотидной последовательности, представленной в последовательностях SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621,

624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696. С другой стороны, изолированная нуклеиновая кислота включает в себя нуклеотидную последовательность, кодирующую полипептид, имеющий идентичность последовательности 80% или более по отношению к аминокислотной последовательности, представленной в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613,

615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697.

С другой стороны, представлены методы идентификации генетического полиморфизма, связанного с изменением уровня биомассы. Методы включают представление популяции растений и определение того, связаны ли с генетической точки зрения один или более генетических полиморфизмов с локусом полипептида, выбранного из группы полипептидов, показанных на фиг. 1-11, и их функциональными гомологами. Производится измерение корреляции между изменением уровня биомассы в тканях растений популяции и наличием одного или более генетических полиморфизмов у растений популяции, что позволяет определить, связаны ли один или более генетических полиморфизмов с таким изменением.

С другой стороны, представлены методы создания линии растений. Методы включают определение того, связаны ли один или более генетических полиморфизмов в популяции растений с локусом одного или нескольких полипептидов, показанных на фиг. 1-11, и функциональными гомологами таких полипептидов. В популяции определяется одно или несколько растений, у которых наличие как минимум одного генетического полиморфизма связано с изменением характеристики биомассы. Описанные выше шаги могут выполняться в любом порядке. Затем одно или несколько из определенных растений самовоспроизводятся или скрещиваются с другим растением для получения семян, как минимум одно потомственное растение, выращенное из такого семени, самовоспроизводится или скрещивается с другим растением. Шаги самоопыления и ауткроссинга повторяются у последующих поколений 0-5 для создания линии растений, у которых присутствует как минимум один полиморфизм. Характеристикой биомассы может служить выход сухого вещества, а в качестве популяции растений могут быть использованы растения проса.

В настоящем документе также описан метод изменения уровня биомассы у растения. Метод включает модификацию модулирующей биомассу эндогенной нуклеиновой кислоты, нуклеиновой кислоты,

включающей нуклеотидную последовательность с открытой рамкой считывания, имеющую идентичность последовательности 80% или более по отношению к нуклеотидной последовательности, выбранной из группы последовательностей SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 и 696. Уровень биомассы растения отличается от соответствующего уровня биомассы контрольного растения, в котором нуклеиновая кислота не модифицирована. Модификацию можно осуществить путем проведения генетической модификации в локусе, содержащем нуклеиновую кислоту. Далее метод может включать отбор растений с измененной биомассой.

В некоторых вариантах эндогенная нуклеиновая кислота кодирует полипептид, имеющий идентичность последовательности 80% или более по отношению к аминокислотной последовательности, выбранной из группы последовательностей SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77,

79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 и 697.

В некоторых вариантах эндогенная нуклеиновая кислота включает в себя нуклеотидную последовательность с открытой рамкой считывания и идентичностью последовательности 90% или более по отношению к нуклеотидной последовательности, выбранной из группы последовательностей SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69,

73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 и 696.

В некоторых вариантах эндогенная нуклеиновая кислота включает в себя нуклеотидную последовательность с открытой рамкой считывания и идентичностью последовательности 95% или более по отношению к нуклеотидной последовательности, выбранной из группы последовательностей SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462,

465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 и 696.

В настоящем документе также описан метод получения растения. Метод включает выращивание растительной клетки, содержащей модифицированную эндогенную нуклеиновую кислоту, кодирующую полипептид, у которого показатель НММ аминокислотной последовательности полипептида больше 65 (показатели НММ получены из аминокислотных последовательностей, показанных на фиг. 1-11), при котором уровень биомассы растения отличается от соответствующего уровня биомассы контрольного растения, в котором нуклеиновая кислота не модифицирована.

С другой стороны, в настоящем документе приведено описание растительной клетки, содержащей модифицированную эндогенную нуклеиновую кислоту, кодирующую полипептид, у которого показатель НММ аминокислотной последовательности полипептида больше 65 (показатели НММ получены из аминокислотных последовательностей, показанных на фиг. 1-11), при котором уровень биомассы растения, полученного из растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором нуклеиновая кислота не модифицирована.

В настоящем документе также приведено описание растительной клетки, содержащей модифицированную эндогенную нуклеиновую кислоту, модулирующую биомассу. Нуклеиновая кислота содержит нуклеотидную последовательность с открытой рамкой считывания и идентичностью последовательности 80% или более по отношению к нуклеотидной последовательности, выбранной из группы последовательностей SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143,

148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 и 696. Уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором нуклеиновая кислота не модифицирована.

У растительной клетки, описанной в настоящем документе, эндогенная нуклеиновая кислота может кодировать полипептид, имеющий идентичность последовательности 80% или более по отношению к аминокислотной последовательности, выбранной из группы последовательностей SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273,

275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 и 697, и при этом уровень биомассы растения, полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором нуклеиновая кислота не модифицирована.

С другой стороны, в настоящем документе приведено описание метода модуляции уровня биомассы у растения. Метод включает интродукцию в растительную клетку экзогенной нуклеиновой кислоты, при этом экзогенная нуклеиновая кислота кодирует полипептид с активным ферментом Е.С. 2.6.1.83.

Также приведено описание растительной клетки, которая содержит экзогенную нуклеиновую кислоту, при этом экзогенная нуклеиновая кислота кодирует полипептид с активным ферментом Е.С. 2.6.1.83, в результате чего уровень биомассы растения полученного из такой растительной клетки, отличается от соответствующего уровня биомассы контрольного растения, в котором отсутствует нуклеиновая кислота.

Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют такое же значение, как его обычно понимают специалисты с обычной квалификацией, работающие в области, к которой относится данное изобретение. Несмотря на то, что при практическом использовании изобретения могут применяться методы и материалы, сходные с теми или аналогичные тем, которые описаны в настоящем документе, ниже приведено описание подходящих методов и материалов. Все публикации, заявки на патент, патенты и прочие ссылочные материалы, приведенные в настоящем документе, включены путем ссылки в полном объеме. В случае противоречия преимущественную силу имеет настоящее описание изобретения, включая определения. Кроме того, материалы, методы и примеры носят исключительно иллюстративный характер и не устанавливают ограничений.

Детали одного или более вариантов осуществления изобретения изложены ниже в сопроводительных чертежах и описании. Другие признаки, цели и преимущества изобретения будут продемонстрированы в описании, чертежах и пунктах формулы изобретения. Слово «включающий» в пунктах формулы изобретения может быть заменено на «состоящий главным образом из» или «состоящий из» согласно общепринятой практике патентного права.

ОПИСАНИЕ ЧЕРТЕЖЕЙ

На фиг. 1 представлено выравнивание аминокислотной последовательности CW00733, соответствующей Ceres Clone: 1384304 (SEQ ID NO: 554) с гомологическими и/или ортологическими аминокислотными последовательностями. На всех фигурах с линейной последовательностью, показанных в настоящем документе, тире в построенной в линию последовательности обозначает делецию, то есть отсутствие аминокислоты в данном месте. Идентичные аминокислоты или консервативные заменители аминокислот среди выровненных последовательностей обозначены прямоугольниками. Фиг. 1 и другие фигуры с линейной последовательностью, приведенные в настоящем документе, созданы с помощью программы MUSCLE версии 3.52.

На фиг. 2 представлено выравнивание аминокислотной последовательности CW00319, соответствующей Ceres Annot: 544549

(SEQ ID NO: 263) с гомологическими и/или ортологическими последовательностями аминокислот.

На фиг. 3 представлено выравнивание аминокислотной последовательности CW00710, соответствующей Ceres Annot: 1355066 (SEQ ID NO: 117) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 4 представлено выравнивание аминокислотной последовательности CW00628, соответствующей антисмысловой последовательности Os01g58420 (SEQ ID NO: 1) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 5 представлено выравнивание аминокислотной последовательности CW00297, соответствующей Ceres Clone: 625057 (SEQ ID NO: 645) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 6 представлено выравнивание аминокислотной последовательности CW00604, соответствующей Ceres Clone: 1356785 (SEQ ID NO: 253) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 7 представлено выравнивание аминокислотной последовательности CW00564, соответствующей Ceres Clone: 638126 (SEQ ID NO: 323) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 8 представлено выравнивание аминокислотной последовательности CW00010, соответствующей Ceres Clone: 26006 (SEQ ID NO: 595) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 9 представлено выравнивание аминокислотной последовательности CW00469, соответствующей Ceres Clone: 4831 (SEQ ID NO: 77) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 10 представлено выравнивание аминокислотной последовательности CW00536, соответствующей Ceres Annot: 847799 (SEQ ID NO: 209) с гомологическими и/или ортологическими аминокислотными последовательностями.

На фиг. 11 представлено выравнивание аминокислотной последовательности CW00191, соответствующей Ceres Annot: 878355

(SEQ ID NO: 426) с гомологическими и/или ортологическими аминокислотными последовательностями.

ПОДРОБНОЕ ОПИСАНИЕ

В изобретении приведено описание методов и материалов, относящихся к модуляции уровня биомассы растений. В некоторых вариантах растения могут также иметь модулированный уровень, например, лигнина, модифицированную архитектуру корневой системы, модифицированную гербицидную устойчивость, модифицированный биосинтез каротеноидов или модулированный состав клеточной оболочки. Методы могут включать трансформацию растительной клетки, при которой нуклеиновая кислота кодирует модулирующий биомассу полипептид, в результате чего экспрессия полипептида приводит к модуляции уровня биомассы. Растительные клетки, продуцированные с использованием таких методов, могут выращиваться для получения растений с увеличенной или уменьшенной биомассой.

Такие растения и семена таких растений могут быть использованы, например, для производства биомассы, имеющей повышенную ценность, для использования в качестве исходного сырья для биотоплива.

I. Термины и определения

«Аминокислота» относится к одной из двадцати аминокислот, встречающихся в природе, и к синтетическим аминокислотам, в том числе оптическим изомерам D- и L-форм.

«Предпочтительный промотор клетки» или «предпочтительный промотор ткани» относится к промотору, который преимущественно стимулирует экспрессию, соответственно, в клетке-мишени или ткани-мишени, однако также может привести к определенной транскрипции в клетках или тканях других типов.

«Контрольное растение» относится к растению, которое не содержит экзогенной нуклеиновой кислоты, присутствующей в рассматриваемом трансгенном растении, однако в остальном имеет такое же или подобное генетическое окружение, что и трансгенное растение. Подходящим контрольным растением может быть нетрансгенное растение дикого типа, нетрансгенный сегрегант, полученный в результате эксперимента по трансформации, или

трансгенное растение, которое содержит экзогенную нуклеиновую кислоту, отличающуюся от рассматриваемой экзогенной нуклеиновой кислоты.

«Домены» – это группы в основном заменимых аминокислот в полипептиде, которые могут быть использованы для описания семейств белков и/или частей белков. Такие домены имеют «отпечаток пальцев» или «сигнатуру», которые могут включать в себя консервативную первичную последовательность, вторичную структуру и/или трехмерную конформацию. Как правило, домены связаны с деятельностью *in vitro* и/или *in vivo*. Длина домена может быть от 10 до 400 аминокислот, например: 10-50 аминокислот, или 25-100 аминокислот, или 35-65 аминокислот, или 35-55 аминокислот, или 45-60 аминокислот, или 200-300 аминокислот, или 300- 400 аминокислот.

«Понижающая регуляция» относится к регуляции, которая снижает продуцирование продуктов экспрессии (мРНК, полипептида или обоих) по отношению к базальному или нативному состоянию.

«Экзогенный» по отношению к нуклеиновой кислоте указывает на то, что нуклеиновая кислота является частью рекомбинантной структуры нуклеиновой кислоты или что она не находится в своей естественной среде. Например, экзогенная нуклеиновая кислота может быть последовательностью от одного вида, интродуцированной в другой вид, то есть гетерологичной нуклеиновой кислотой. В основном такую экзогенную нуклеиновую кислоту интродуцируют в другие виды через рекомбинантную конструкцию нуклеиновой кислоты. Экзогенная нуклеиновая кислота может также быть последовательностью, которая является нативной для какого-либо организма и которая реинтродуцируется в клетки этого организма. Экзогенную нуклеиновую кислоту, содержащую нативную последовательность, можно часто отличить от последовательности, образовавшейся естественным путем, по наличию неприродных последовательностей, присоединенных к экзогенной нуклеиновой кислоте, например ненативных регуляторных последовательностей, примыкающих к нативной последовательности в рекомбинантной структуре нуклеиновой кислоты. Кроме того, устойчиво трансформированные экзогенные нуклеиновые кислоты обычно

интегрированы в места, которые отличаются от тех мест, где находится нативная последовательность. Следует иметь в виду, что экзогенные нуклеиновые кислоты могут быть интродуцированы в клетку-предшественник, а не в рассматриваемую клетку. Например, трансгенное растение, содержащее экзогенную нуклеиновую кислоту, может быть потомством от скрещивания между устойчиво трансформированным растением и нетрансгенным растением. Считается, что такое потомство содержит экзогенную нуклеиновую кислоту.

«Экспрессия» относится к процессу преобразования генетической информации полинуклеотида в РНК путем транскрипции, которая катализируется ферментом полимеразой РНК, и в белок путем трансляции мРНК на рибосомах.

«Гетерологичный полипептид» для целей настоящего документа относится к полипептиду, который не является естественным для растительной клетки полипептидом, например трансгенное растение *Panicum virgatum*, трансформированное с помощью кодирующей последовательности и экспрессирующее кодирующую последовательность для полипептида, обеспечивающего транспорт азота из растения *Zea mays*.

«Изолированная нуклеиновая кислота» для целей настоящего документа включает природную нуклеиновую кислоту, при условии что одна или обе последовательности, непосредственно примыкающие к данной нуклеиновой кислоте в ее природном геноме, удалены или отсутствуют. Таким образом, изолированная нуклеиновая кислота включает, помимо прочего, нуклеиновую кислоту, которая существует в виде очищенной молекулы или молекулы нуклеиновой кислоты, включенной в вектор или вирус. Нуклеиновая кислота, существующая среди сотен и миллионов других нуклеиновых кислот, например, в библиотеках кДНК, геномных библиотеках или телевых срезах, содержащих фрагменты рестрикции геномной ДНК, не должна рассматриваться как изолированная нуклеиновая кислота.

«Модуляция» уровня биомассы относится к изменению уровня биомассы, которое наблюдается в результате экспрессии или транскрипции экзогенной нуклеиновой кислоты в растительной клетке и/или растении. Изменение уровня измеряется относительно

соответствующего уровня у контрольных растений.

«Нуклеиновая кислота» и «полинуклеотид» в настоящем документе являются взаимозаменяемыми и относятся как к РНК, так и к ДНК, в том числе кДНК, геномной ДНК, синтетической ДНК и ДНК или РНК, содержащим аналоги нуклеиновых кислот. Нуклеиновая кислота может быть двухцепочечной или одноцепочечной (то есть кодирующей цепью или антисмысловой цепью). Неограниченные примеры полинуклеотидов включают гены, фрагменты генов, экзоны, интроны, матричную РНК (мРНК), транспортную РНК, рибосомальную РНК, синтетическую РНК, микроРНК, рибозимы, кДНК, рекомбинантные полинуклеотиды, разветвленные полинуклеотиды, зонды нуклеиновой кислоты и праймеры нуклеиновой кислоты. Полинуклеотид может содержать нестандартные или модифицированные нуклеотиды.

«Функционально связанный» относится к размещению регуляторной области и последовательности, транскрибируемой в нуклеиновой кислоте, с тем чтобы регуляторная область была эффективной для регулирования транскрипции или трансляции последовательности. Например, чтобы функционально связать кодирующую последовательность и регуляторную область, сайт инициации трансляции трансляционной рамки считывания кодирующей последовательности обычно располагают между первым и пятидесятым нуклеотидом нижележащей регуляторной области. Однако регуляторная область может быть расположена примерно на 5000 нуклеотидов выше сайта инициации трансляции или примерно на 2000 нуклеотидов выше сайта инициации транскрипции.

«Полипептид» для целей настоящего документа относится к соединению из двух или более субъединичных аминокислот, аналогов аминокислот или других пептидомиметиков, независимо от посттрансляционных модификаций, например фосфорилирования или гликозилирования. Субъединицы могут быть связаны пептидными связями или другими связями, такими как, например, сложноэфирные или простые эфирные связи. Данное определение охватывает полноразмерные полипептиды, усеченные полипептиды, точечные мутанты, инсерционные мутанты, сплайс-варианты, химерные белки и их фрагменты.

«Потомство» включает потомков определенного растения или

линии растений. Потомство скороспелого растения включает семена, сформированные на растениях поколений F1, F2, F3, F4, F5, F6 и последующих поколений, или семена, сформированные на растениях поколений BC1, BC2, BC3 и последующих поколений, или семена, сформированные на растениях поколений F1BC1, F1BC2, F1BC3 и последующих поколений. Обозначение F1 относится к потомству от скрещивания двух родительских организмов с четко выраженными генетическими признаками. Обозначения F2, F3, F4, F5 и F6 относятся к последующим поколениям потомков растения F1, полученным в результате самоопыления или переопыления.

«Регуляторная область» относится к нуклеиновой кислоте, имеющей нуклеотидные последовательности, которые оказывают влияние на инициацию и скорость транскрипции или трансляции, а также на стабильность и/или подвижность продукта транскрипции или трансляции. Регуляторные области включают, помимо прочего, последовательности промотора, последовательности генов-усилителей, ответные элементы, участки распознавания белка, индуцируемые элементы, белок-связывающие последовательности, нетранслируемые области 5' и 3', сайты инициации транскрипции, терминирующие последовательности, последовательности полиаденилирования, интроны и их сочетания. Регуляторная область обычно включает как минимум коровый (базальный) промотор. Регуляторная область также может включать как минимум один управляющий элемент, такой как последовательность генов-усилителей, вышележащий элемент или вышележащая область активации. Например, подходящим геном-усилителем является цис-действующий регуляторный элемент (от -212 до -154) от вышележащей области гена октопин-синтазы (ocs). Fromm et al., *The Plant Cell*, 1:977-984 (1989).

«Повышающая регуляция» относится к регуляции, которая повышает уровень продукта экспрессии (мРНК, полипептида или обоих) по отношению к базальному или нативному состоянию.

«Вектор» относится к репликону, такому как плазмид, фаг или космида, в который может быть введен другой сегмент ДНК, чтобы добиться репликации введенного сегмента. Как правило, вектор способен к репликации, когда он связан с соответствующими

управляющими элементами. Термин «вектор» включают клонирующий и экспрессионный векторы, а также вирусные векторы и интегрирующие векторы. «Экспрессионный вектор» – это вектор, который включает регуляторную область.

II. Полипептиды

Полипептиды, описанные в настоящем документе, включают модулирующие биомассу полипептиды. Модулирующие биомассу полипептиды могут быть эффективными для модуляции уровней биомассы при экспрессировании в растении или растительной клетке. Такие полипептиды обычно содержат как минимум один домен, указывающий на модулирующие биомассу полипептиды, как подробно описано в настоящем документе. Модулирующие биомассу полипептиды обычно имеют показатель HMM более 65, как подробно описано в настоящем документе. В некоторых вариантах модулирующие биомассу полипептиды более чем на 80% идентичны последовательностям SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440,

442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697, как более подробно описано в настоящем документе.

А. Домены, указывающие на модулирующие биомассу полипептиды

Модулирующий биомассу полипептид может содержать ДНК-связывающий домен «цинковые пальцы» (zf-Dof), который, предположительно, характерен для модулирующего биомассу полипептида. SEQ ID NO: 263 устанавливает аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Annot: 544549 (SEQ ID NO: 262), которая, предположительно, кодирует полипептид, содержащий ДНК-связывающий домен «цинковые пальцы». Например, модулирующий биомассу полипептид может содержать ДНК-связывающий домен «цинковые пальцы», имеющий идентичность последовательности 60% или более по отношению к остаткам 130-192 последовательности SEQ ID NO: 263. В некоторых вариантах модулирующий биомассу полипептид может содержать ДНК-связывающий домен «цинковые пальцы», имеющий идентичность последовательности 60% или более по отношению к ДНК-связывающему домену «цинковые пальцы» одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320 или 321. ДНК-связывающие домены «цинковые пальцы» таких последовательностей приводятся в

перечне последовательностей. Домены «цинковые пальцы» (Znf) представляют собой относительно небольшие белковые мотивы, которые связывают один или несколько атомов цинка и которые обычно содержат несколько пальцевидных ответвлений, которые устанавливают контакт со своей молекулой-мишенью. Ранее они обозначались как ДНК-связывающий мотив в факторе транскрипции TFIIIA из *Xenopus laevis*, однако в настоящее время определено, что они связывают ДНК, РНК, белок и/или липидные субстраты. Их связывающие свойства зависят от аминокислотной последовательности доменов «цинковые пальцы» и от линкера между «пальцами», а также от структур более высокого порядка и количества «пальцев». Домены «цинковые пальцы» (Znf) часто находятся в кластерах, у которых «пальцы» имеют различную связывающую специфику. Существует множество суперсемейств мотивов Znf, отличающихся как по последовательности, так и по структуре. Они демонстрируют многообразие режимов связывания (например, некоторые связывают ДНК, другие – белок), что указывает на то, что мотивы Znf являются стабильными структурными каркасами, которые развили специализированные функции. Например, функцию Znf-содержащих белков в транскрипции и трансляции генов, направленный перенос мРНК, формирование клеточного скелета, развитие эпителиальных тканей, клеточную адгезию, укладку белка, реконструкцию хроматина, распознавание цинка и многие другие. Цинк-связывающие мотивы являются стабильными структурами и редко подвергаются конформационным изменениям при связывании своей мишени. Ортологи DOF 1.3 могут содержать ДНК-связывающие домены «цинковые пальцы».

Модулирующий биомассу полипептид может содержать домен фитохелатин-синтетаза, который, предположительно, характерен для модулирующего биомассу полипептида. SEQ ID NO: 117 устанавливает аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Annot: 1355066 (SEQ ID NO: 116), которая, предположительно, кодирует полипептид, содержащий домен фитохелатин-синтетаза. Например, модулирующий биомассу полипептид может содержать домен фитохелатин-синтетаза, имеющий идентичность последовательности

60% или более по отношению к остаткам 44–208 последовательности SEQ ID NO: 117. В некоторых вариантах модулирующий биомассу полипептид может содержать область с доменом фитохелатин-синтетаза, имеющим идентичность последовательности 60% или более по отношению к области с доменом фитохелатин-синтетаза одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206 или 207. Области с доменом фитохелатин-синтетаза таких последовательностей устанавливаются в перечне последовательностей. Белок фитохелатин-синтетаза может быть ферментом, ответственным за синтез связывающих тяжелые металлы пептидов (фитохелатинов) из глутатиона и родственных тиолов. Фермент обычно катализирует деглицинацию молекулы-донора глутатиона (GSH). Фермент обычно содержит каталитическую триаду остатков цистеина, гистидина и аспартата.

Модулирующий биомассу полипептид может содержать домен AP2, который, предположительно, характерен для модулирующего биомассу полипептида. SEQ ID NO: 1 устанавливает аминокислотную последовательность клона *Oryza sativa*, обозначенного в настоящем документе как Os01g58420, который, предположительно, кодирует полипептид, содержащий домен AP2. Например, модулирующий биомассу полипептид может содержать домен AP2 с идентичностью последовательности 60% или более по отношению к остаткам 32–83 последовательности SEQ ID NO: 1. В некоторых вариантах модулирующий биомассу полипептид может содержать домен AP2 с идентичностью последовательности 60% или более по отношению к домену AP2 одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74 или 75. Домены AP2 таких последовательностей устанавливаются в перечне

последовательностей. В некоторых вариантах в растении экспрессируется антисмысловая последовательность, которая модулирует биомассу, как описано в настоящем документе. Например, для модуляции биомассы в растении может быть экспрессирована антисмысловая последовательность нуклеиновой кислоты Os01g58420, такая как последовательность SEQ ID NO: 678. Остатки аминокислоты домена AP2 могут связываться с ДНК и обычно находятся в белках-факторах транскрипции.

Модулирующий биомассу полипептид может содержать домен аминотрансферазы класса I и II, который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 645 устанавливает аминокислотную последовательность клона *Glycine max*, обозначенного в настоящем документе как Ceres Clone: 625057 (SEQ ID NO: 644), который, предположительно, кодирует полипептид, содержащий домен аминотрансферазы класса I и II. Например, модулирующий биомассу полипептид может содержать домен аминотрансферазы класса I и II, имеющий идентичность последовательности 60% или более по отношению к остаткам 88-453 последовательности SEQ ID NO: 645. В некоторых вариантах модулирующий биомассу полипептид может содержать домен аминотрансферазы класса I и II, имеющий идентичность последовательности 60% или более по отношению к домену аминотрансферазы класса I и II одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677 или 689. Домены аминотрансферазы класса I и II таких последовательностей приведены в перечне последовательностей. Аминотрансферазы имеют определенные конструктивные особенности, общие с другими пиридоксальфосфат-зависимыми энзимами, например ковалентное связывание пиридоксальфосфатной группы с остатками лизина. На основе сходства последовательностей эти разнообразные ферменты можно сгруппировать в класс I и класс II. Примерами полипептидов, содержащих домены аминотрансферазы класса I и II, могут быть полипептиды LL-DAP (EC 2.6.1.83) (Watanabe et al., Mechanism of Substrate Recognition and PLP-induced

Conformational Changes in LL-Diaminopimelate aminotransferase from *Arabidopsis thaliana*. *J. Mol. Biol.* 384, 1314-1329 (2008)). Полипептид LL-DAP катализирует взаимное превращение LL-2,6-диаминогептандиоата и 2-оксоглутарата в (S)-2,3,4,5-тетрагидропиридин-2,6-дикарбоксилат, L-глутамат и воду.

Модулирующий биомассу полипептид может содержать ДНК-связывающий домен семейства Myb, который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 323 представляет аминокислотную последовательность клона *Glycine max*, обозначенного в настоящем документе как Ceres Clone: 638126 (SEQ ID NO: 321), который, предположительно, кодирует полипептид, содержащий ДНК-связывающий домен семейства Myb. Например, модулирующий биомассу полипептид может содержать ДНК-связывающий домен семейства Myb, имеющий идентичность последовательности 60% или более по отношению к остаткам 13-62 последовательности SEQ ID NO: 323. В некоторых вариантах модулирующий биомассу полипептид может содержать ДНК-связывающий домен семейства Myb, имеющий идентичность последовательности 60% или более по отношению к ДНК-связывающему домену семейства Myb одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422 или 424. ДНК-связывающие домены семейства Myb таких последовательностей представлены в перечне последовательностей. Семейство ДНК-связывающих доменов Myb включает в себя ДНК-связывающие домены из белков семейства Myb, а также семейство доменов SANT.

Модулирующий биомассу полипептид может содержать домен альфа/бета-складки гидролазы, который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 595 представляет аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Clone: 26006 (SEQ ID NO: 594), который,

предположительно, кодирует полипептид, содержащий домен альфа/бета-складки гидролазы. Например, модулирующий биомассу полипептид может содержать домен альфа/бета-складки гидролазы, имеющий идентичность последовательности 60% или более по отношению к остаткам 35-257 последовательности SEQ ID NO: 595. В некоторых вариантах модулирующий биомассу полипептид может содержать область с доменом альфа/бета-складки гидролазы, имеющий идентичность последовательности 60% или более по отношению к домену альфа/бета-складки гидролазы одного или нескольких полипептидов, представленных в последовательностях SEQ ID NO: 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643 или 691. Домены альфа/бета-скачков гидролазы таких последовательностей приведены в перечне последовательностей. Альфа/бета-складка гидролазы является общей для ряда гидролитических ферментов, имеющих самое разное филогенетическое происхождение и выполняющих каталитическую функцию. Сердцевиной каждого фермента является альфа/бета слой (нежели ствол), содержащий 8 цепей, соединенных спиралями. Считается, что ферменты произошли от общего предшественника, сохранив при этом строение каталитических остатков. Все они имеют каталитическую триаду, элементы которой зарождаются на петлях, являющихся лучшими консервативными структурными особенностями складки.

Модулирующий биомассу полипептид может содержать домен «фактор быстрого подщелачивания» (ФБП), который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 77 представляет аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Clone:4831 (SEQ ID NO: 76), который, предположительно, кодирует полипептид, содержащий домен ФБП. Например, модулирующий биомассу полипептид может содержать домен ФБП, имеющий идентичность последовательности 60% или более по отношению к остаткам 57-129 последовательности SEQ ID NO: 77. В некоторых вариантах модулирующий биомассу полипептид может содержать домен ФБП,

имеющий идентичность последовательности 60% или более по отношению к домену ФБП одного или более полипептидов, представленных в последовательностях SEQ ID NO: 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114 или 115. Домены ФБП таких последовательностей приведены в перечне последовательностей. Домены ФБП обычно находятся в распространенных пептидах 5-kDa в растениях, которые, как сообщается, играют роль в подавлении роста и развития корня у некоторых растений.

Модулирующий биомассу полипептид может содержать домен DUF640, который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 209 приводит аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Annot: 847799 (SEQ ID NO: 208), который, предположительно, кодирует полипептид, содержащий домен DUF640. Например, модулирующий биомассу полипептид может содержать домен DUF640, имеющий идентичность последовательности 60% или более по отношению к остаткам 19-152 последовательности SEQ ID NO: 209. В некоторых вариантах модулирующий биомассу полипептид может содержать домен DUF640, имеющий идентичность последовательности 60% или более по отношению к домену DUF640 одного или нескольких полипептидов, приведенных в последовательностях SEQ ID NO: 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250 или 251. Домены DUF640 таких последовательностей представлены в перечне последовательностей.

Модулирующий биомассу полипептид может содержать домен семейства PTR2 POT, который, предположительно, характерен для модулирующего биомассу полипептида. Последовательность SEQ ID NO: 426 представляет аминокислотную последовательность клона *Arabidopsis*, обозначенного в настоящем документе как Ceres Annot: 878355 (SEQ ID NO: 425), который, предположительно, кодирует полипептид, содержащий домен семейства PTR2 POT. Например, модулирующий биомассу полипептид может содержать домен семейства PTR2 POT, имеющий идентичность последовательности 60%

или более по отношению к остаткам 100–509 последовательности SEQ ID NO: 426. В некоторых вариантах модулирующий биомассу полипептид может содержать домен семейства PTR2 POT, имеющий идентичность последовательности 60% или более по отношению к домену семейства PTR2 POT одного или более полипептидов, представленных в последовательностях SEQ ID NO: 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 693, 695 или 697. Области с доменом семейства PTR2 POT таких последовательностей приведены в перечне последовательностей. Транспорт пептидов в клетки является биологическим феноменом, убедительно подтвержденным документальными доказательствами. Он осуществляется специфическими энергозависимыми переносчиками и присутствует в самых разнообразных организмах – от бактерий до человека. Семейство белков PTR отличается от ABC-транспортёров пептидов и было открыто с помощью секвенирования ряда недавно открытых транспортных белков для переноса пептидов. Представляется, что данные белки главным образом участвуют во введении небольших пептидов с сопутствующим поглощением протона. В некоторых вариантах белок POT, как описано в настоящем документе, может включать N-концевой сигнальный пептид. В некоторых вариантах сигнальный пептид может быть характерным для плазматической мембраны. В некоторых вариантах сигнальный пептид может быть характерным для мембраны эндоплазматической сети или хлоропластовой мембраны. Примеры сигнальных пептидов показаны в заявке, в перечне последовательностей. Для прогнозирования наличия и типа транзитных пептидов могут применяться методы биоинформатики. Такие методы не опираются исключительно на сходство последовательности. Поскольку ортологические белки чаще имеют одинаковую локализацию, степень сходства последовательности, необходимая для того, чтобы сделать вывод о

со-локализации, выше, чем у аналогичной трехмерной структуры, а изоформы одного и того же белка могут иметь различную локализацию. Для прогнозирования сигнальных пептидов можно использовать программу для прогнозирования внутриклеточной локализации белка WoLF PSORT (Horton et al., 2007 "WoLF PSORT: Protein Localization Predictor", Nucleic Acids Research, doi:10.1093/nar/gkm259, 2007; Horton et al., 2006 "Protein Subcellular Localization Prediction with WoLF PSORT", Proceedings of the 4th Annual Asia Pacific Bioinformatics Conference APBC06, Taipei, Taiwan. pp. 39-48, 2006). Примеры сигнальных пептидов можно получить из последовательностей, имеющих в открытом доступе, с помощью анализа последовательности, выполненного программой WoLF PSORT, которая предоставляет данные о многочисленных ортологических сигнальных пептидах.

В эукариотных организмах присутствует несколько типов сигнальных пептидов и связанных с ними сортирующих сигналов, каждый из которых участвует в мембранной транслокации и/или инсерции. Как правило, сигнальные пептиды, специфические для эндоплазматического ретикулума (ER), являются котрансляционными, в то время как сигнальные пептиды, специфические для митохондрий или хлоропластов, являются посттрансляционными, но не развернутыми шаперонами. Например, N-концевой сигнал с гидрофобной секцией переменной длины вызывает котрансляционный переход белков через или в мембрану эндоплазматического ретикулума. N-концевые сигналы большей частью не зависят от белков-носителей. Такие сигнальные пептиды обычно являются взаимозаменяемыми с различными белками, обычно расщеплены и обычно ограничиваются первыми приблизительно 90 аминокислотными остатками. Расщепление на N-конце и котрансляционное распознавание делают сигнальные пептиды типично ортогональными по отношению к функции белка, однако это является только общим сходством. В некоторых вариантах белок POT, как описано в настоящем документе, может включать C-концевой сортирующий сигнал. Примеры C-концевых сортирующих сигналов включают, в числе прочего, сигнал KDEL (растворение) или KKXX (мембранный

белок) для удержания ER, SKL для пероксисомального воздействия (растворение), NPIR для вакуоли и LPXTG для стенки бактериальной клетки. В некоторых вариантах белок POT, как описано в настоящем документе, может включать внутренний сортирующий сигнал. Такие сигналы включают клеточные сигналы внутриядерной локализации, которые наблюдаются на поверхности белка со складчатой структурой, однако могут находиться в любом месте одноразмерной последовательности. В некоторых вариантах белок POT, как описано в настоящем документе, может включать N-концевой сигнальный пептид длиной примерно 95, 90, 85, 80, 75, 70, 65, 60, 55, 50, 45, 40, 35, 30, 25, 20, 15, 10 или 5 аминокислот, начинающийся с N-конца указанного белка POT. В некоторых вариантах в белке POT, как описано в настоящем документе, может отсутствовать весь N-концевой сигнальный пептид или его часть. В некоторых вариантах в белке POT, как описано в настоящем документе, N-концевой сигнальный пептид может быть удален и заменен на другой N-концевой сигнальный пептид. Например, специалист в данной области может удалить или синтезировать последовательность без 45 аминокислот с N-концом последовательности SEQ ID NO: 426 и добавить методом слияния или синтеза другой сигнальный пептид, специфичный для такой же мембраны или другой мембраны мишени.

В некоторых вариантах модулирующий биомассу полипептид усекается в аминоконцевой или карбоксиконцевой области природного полипептида. Усеченный полипептид может удерживать определенные домены природного полипептида при отсутствии других. Таким образом, при вариантах, когда длина меньше или больше на 5 аминокислот, обычно показывают активность усеченного полипептида по модуляции биомассы. В некоторых вариантах усеченный полипептид является доминантным негативным полипептидом. Экспрессия в растении с таким усеченным полипептидом придает различие в уровне биомассы растения по сравнению с соответствующим уровнем биомассы контрольного растения, у которого нет усеченного полипептида.

Б. Функциональные гомологи, определяемые с помощью программы Reciprocal BLAST

В некоторых вариантах один или несколько функциональных

гомологов эталонного модулирующего биомассу полипептида, определенного в одном или нескольких указанных выше описаниях базы данных по семействам белков (Pfam), пригодны для использования в качестве модулирующих биомассу полипептидов. Функциональный гомолог является полипептидом, который имеет сходство последовательности с эталонным полипептидом и который выполняет одну или несколько биохимических или физиологических функций эталонного полипептида. Функциональный гомолог и эталонный полипептид могут быть природными полипептидами, а сходство последовательности может быть вызвано конвергентными или дивергентными эволюционными событиями. В силу этого функциональные гомологи иногда обозначают в литературе как гомологи, или ортологи, или паралоги. Варианты природных функциональных гомологов, таких как полипептиды, закодированные мутантами кодирующей последовательности дикого типа, могут сами быть функциональными гомологами. Функциональные гомологи могут быть также созданы сайт-направленным мутагенезом кодирующей последовательности модулирующего биомассу полипептида или путем комбинирования доменов из кодирующих последовательностей для различных природных модулирующих биомассу полипептидов («перестановка доменов»). Термин «функциональный гомолог» иногда применяется к нуклеиновой кислоте, которая кодирует функционально гомологичный полипептид.

Функциональные гомологи можно идентифицировать путем анализа выравнивания последовательностей нуклеотидов и полипептидов. Например, при выполнении поиска в базе данных последовательностей нуклеотидов и полипептидов можно определить гомологи модулирующих биомассу полипептидов. Анализ последовательностей может включать анализ всего банка данных с помощью программы BLAST, Reciprocal BLAST или PSIBLAST с использованием аминокислотной последовательности модулирующего биомассу полипептида в качестве эталонной последовательности. В некоторых случаях аминокислотная последовательность выводится из нуклеотидной последовательности. Те полипептиды в базе данных, которые имеют идентичность последовательности более 40%, являются кандидатами для дальнейшей оценки их пригодности для

использования в качестве модулирующих биомассу полипептидов. Сходство аминокислотной последовательности допускает замену консервативных аминокислот, например замену одного гидрофобного остатка на другой или замену одного полярного остатка на другой. При желании можно выполнить проверку таких кандидатов вручную, чтобы сузить их количество для дальнейшей оценки. При проверке вручную можно выбрать тех кандидатов, у которых в модулирующих биомассу полипептидах, по-видимому, присутствуют домены, например консервативные функциональные домены.

Консервативные области можно определить путем локализации области в пределах первичной аминокислотной последовательности модулирующего биомассу полипептида, то есть повторяющаяся последовательность образует некую вторичную структуру (например, спирали и складчатые бета-слои), устанавливает положительно или отрицательно заряженные домены или представляет собой белковый мотив или домен. Например, см. в Интернете веб-сайт базы данных по семействам белков (Pfam) с описанием консенсусных последовательностей для различных белковых мотивов и доменов на странице sanger.ac.uk/Software/Pfam/ и pfam.janelia.org/. Описание информации, включенной в базу данных Pfam, приведено в работах Sonnhammer *et al.*, *Nucl. Acids Res.*, 26:320-322 (1998); Sonnhammer *et al.*, *Proteins*, 28:405-420 (1997); и Bateman *et al.*, *Nucl. Acids Res.*, 27:260-262 (1999). Консервативные области также можно определить путем выравнивания последовательностей одинаковых или родственных полипептидов от близкородственных видов. Близкородственные виды преимущественно относятся к одному семейству. В некоторых вариантах будет достаточным выравнивание последовательностей от двух разных видов.

Как правило, для определения консервативных областей пригодны полипептиды, показывающие идентичность аминокислотных последовательностей не менее 40%. Консервативные области родственных полипептидов показывают идентичность аминокислотных последовательностей не менее 45% (например, идентичность аминокислотных последовательностей не менее 50%, не менее 60%, не менее 70%, не менее 80% или не менее 90%). В некоторых вариантах консервативная область показывает идентичность

аминокислотных последовательностей не менее 92%, 94%, 96%, 98% или 99%.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 554, приведены на фиг. 1 и в перечне последовательностей. Такие функциональные гомологи содержат, например, CeresAnnot:564098 (SEQ ID NO: 556), CeresAnnot:1443290 (SEQ ID NO: 558), CeresClone:1042157 (SEQ ID NO: 560), CeresClone:1919714 (SEQ ID NO: 562), GI:157336039 (SEQ ID NO: 563), CeresAnnot:8454153 (SEQ ID NO: 565), CeresAnnot:1722302 (SEQ ID NO: 567), CeresAnnot:8733140 (SEQ ID NO: 569), CeresAnnot:1452096 (SEQ ID NO: 571), CeresClone:1645639 (SEQ ID NO: 573), GI:157344920 (SEQ ID NO: 574), GI:115440865 (SEQ ID NO: 575), CeresClone:340925 (SEQ ID NO: 577), CeresAnnot:8669404 (SEQ ID NO: 579), CeresClone:100028078 (SEQ ID NO: 581), CeresAnnot:1503869 (SEQ ID NO: 583), CeresAnnot:1525651 (SEQ ID NO: 585), CeresClone:2031281 (SEQ ID NO: 587), CeresClone:483742 (SEQ ID NO: 589), CeresClone:100802111 (SEQ ID NO: 591), CeresClone:1460255 (SEQ ID NO: 593). В некоторых случаях функциональный гомолог SEQ ID NO: 554 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 554.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 263, приведены на фиг. 2 и в перечне последовательностей. Такие функциональные гомологи содержат, например, GI:157355009 (SEQ ID NO: 264), CeresAnnot:1464457 (SEQ ID NO: 266), CeresClone:1584660 (SEQ ID NO: 268), GI:115474149 (SEQ ID NO: 269), CeresAnnot:8636233 (SEQ ID NO: 271), CeresClone:1777035 (SEQ ID NO: 273), CeresClone:1990929 (SEQ ID NO: 275), GI:194692166 (SEQ ID NO: 276), CeresAnnot:1458507 (SEQ ID NO: 278), GI:147780712 (SEQ ID NO: 279), CeresAnnot:8642924 (SEQ ID NO: 281), GI:115451001 (SEQ ID NO: 282), AAF87041 (SEQ ID NO:

283), CeresClone:1573856 (SEQ ID NO: 285), CeresAnnot:1476818 (SEQ ID NO: 287), CeresAnnot:1450024 (SEQ ID NO: 289), CeresAnnot:1503065 (SEQ ID NO: 291), GI:147866358 (SEQ ID NO: 292), CeresClone:230073 (SEQ ID NO: 294), (SEQ ID NO: 295), (SEQ ID NO: 296), GI:78708599 (SEQ ID NO: 297), GI:15451553 (SEQ ID NO: 298), GI:125542572 (SEQ ID NO: 299), GI:157342426 (SEQ ID NO: 300), CeresAnnot:538622 (SEQ ID NO: 302), CeresAnnot:8460661 (SEQ ID NO: 304), GI:15983797 (SEQ ID NO: 305), GI:115435804 (SEQ ID NO: 306), CeresClone:1599579 (SEQ ID NO: 308), CeresAnnot:1469831 (SEQ ID NO: 310), GI:9758342 (SEQ ID NO: 311), GI:21536859 (SEQ ID NO: 312), CeresClone:113639 (SEQ ID NO: 314), GI:15232818 (SEQ ID NO: 315), CeresClone:1571328 (SEQ ID NO: 317), CeresClone:1868988 (SEQ ID NO: 319), GI:1669341 (SEQ ID NO: 320) или GI:157359317 (SEQ ID NO: 321). В некоторых случаях функциональный гомолог SEQ ID NO: 263 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 263.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 117, приведены на фиг. 3 и в перечне последовательностей. Такие функциональные гомологи включают, например, GI:90657534 (SEQ ID NO: 118), CeresClone:1237946 (SEQ ID NO: 120), GI:118488472 (SEQ ID NO: 121), GI:38194917 (SEQ ID NO: 122), GI:157341292 (SEQ ID NO: 123), CeresClone:1957107 (SEQ ID NO: 125), CeresAnnot:8640603 (SEQ ID NO: 127), CeresClone:829440 (SEQ ID NO: 129), CeresClone:285169 (SEQ ID NO: 131), GI:116790012 (SEQ ID NO: 132), GI:157356290 (SEQ ID NO: 133), CeresAnnot:1450186 (SEQ ID NO: 135), CeresClone:1804732 (SEQ ID NO: 137), CeresClone:1781794 (SEQ ID NO: 139), CeresAnnot:8656625 (SEQ ID NO: 141), GI:162462515 (SEQ ID NO: 142), CeresClone:570485 (SEQ ID NO: 144), GI:125586664 (SEQ ID NO: 145), GI:116788824 (SEQ ID NO: 146), GI:115453531 (SEQ ID

NO: 147), CeresClone:17250 (SEQ ID NO: 149), CeresAnnot:1363625 (SEQ ID NO: 151), GI:75133694 (SEQ ID NO: 152), GI:147780878 (SEQ ID NO: 153), GI:157341291 (SEQ ID NO: 154), GI:38194916 (SEQ ID NO: 155), GI:157356291 (SEQ ID NO: 156), CeresClone:1883580 (SEQ ID NO: 158), CeresClone:1848658 (SEQ ID NO: 160), CeresAnnot:1450185 (SEQ ID NO: 162), GI:13477083 (SEQ ID NO: 163), GI:115463639 (SEQ ID NO: 164), CeresClone:98007 (SEQ ID NO: 166), CeresAnnot:1326475 (SEQ ID NO: 168), GI:115473243 (SEQ ID NO: 169), CeresAnnot:870466 (SEQ ID NO: 171), CeresClone:1806851 (SEQ ID NO: 173), GI:75133695 (SEQ ID NO: 174), CeresClone:1788775 (SEQ ID NO: 176), CeresClone:1546455 (SEQ ID NO: 178), CeresClone:1902642 (SEQ ID NO: 180), CeresAnnot:8632643 (SEQ ID NO: 182), CeresClone:236876 (SEQ ID NO: 184), GI:90657629 (SEQ ID NO: 185), GI:30090032 (SEQ ID NO: 186), CeresAnnot:8640602 (SEQ ID NO: 188), GI:115453533 (SEQ ID NO: 189), GI:162462330 (SEQ ID NO: 190), GI:38230578 (SEQ ID NO: 191), CeresAnnot:8632641 (SEQ ID NO: 193), GI:168016456 (SEQ ID NO: 194), GI:125532513 (SEQ ID NO: 195), GI:157354382 (SEQ ID NO: 196), CeresAnnot:1481980 (SEQ ID NO: 198), CeresAnnot:1535466 (SEQ ID NO: 200), CeresAnnot:1297618 (SEQ ID NO: 202), GI:119040466 (SEQ ID NO: 203), GI:116310381 (SEQ ID NO: 204), CeresAnnot:8702104 (SEQ ID NO: 206) или GI:157340500 (SEQ ID NO: 207). В некоторых случаях функциональный гомолог SEQ ID NO: 117 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 117.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 1, приведены на фиг. 4 и в перечне последовательностей. Такие функциональные гомологи включают, например, GI:84795244 (SEQ ID NO: 2), CeresClone:1725396 (SEQ ID NO: 4), CeresAnnot:8669118 (SEQ ID NO: 6), CeresClone:280241 (SEQ ID NO: 8), CeresClone:1712594 (SEQ ID NO: 10), GI:190361125 (SEQ ID NO:

11), GI:4099921 (SEQ ID NO: 12), GI:147844573 (SEQ ID NO: 13), GI:67906426 (SEQ ID NO: 14), GI:57012757 (SEQ ID NO: 15), GI:56567583 (SEQ ID NO: 16), GI:84795246 (SEQ ID NO: 17), GI:84795248 (SEQ ID NO: 18), CeresClone:1805203 (SEQ ID NO: 20), CeresClone:101497672 (SEQ ID NO: 22), CeresClone:224845 (SEQ ID NO: 24), GI:115464685 (SEQ ID NO: 25), CeresClone:1287030 (SEQ ID NO: 27), CeresAnnot:8733383 (SEQ ID NO: 29), GI:84795240 (SEQ ID NO: 30), CeresClone:1806017 (SEQ ID NO: 32), GI:84795242 (SEQ ID NO: 33), GI:84795238 (SEQ ID NO: 34), CeresClone:1733772 (SEQ ID NO: 36), GI:37625037 (SEQ ID NO: 37), GI:37625035 (SEQ ID NO: 38), GI:147805535 (SEQ ID NO: 39), GI:157358724 (SEQ ID NO: 40), GI:4099914 (SEQ ID NO: 41), CeresAnnot:1520029 (SEQ ID NO: 43), CeresClone:1065091 (SEQ ID NO: 45), CeresClone:1793792 (SEQ ID NO: 47), CeresClone:1619220 (SEQ ID NO: 49), GI:57012875 (SEQ ID NO: 50), GI:147811787 (SEQ ID NO: 51), CeresClone:1842925 (SEQ ID NO: 53), GI:20340233 (SEQ ID NO: 54), CeresClone:1657843 (SEQ ID NO: 56), CeresAnnot:1455887 (SEQ ID NO: 58), GI:118490009 (SEQ ID NO: 59), CeresClone:1381515 (SEQ ID NO: 61), CeresClone:22775 (SEQ ID NO: 63), GI:60459377 (SEQ ID NO: 64), CeresAnnot:1488231 (SEQ ID NO: 66), CeresClone:1884969 (SEQ ID NO: 68), CeresClone:1802100 (SEQ ID NO: 70), GI:156145802 (SEQ ID NO: 71), GI:28274832 (SEQ ID NO: 72), CeresClone:568399 (SEQ ID NO: 74) или GI:115460458 (SEQ ID NO: 75). В некоторых случаях функциональный гомолог SEQ ID NO: 1 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 1.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 645, приведены на фиг. 5 и в перечне последовательностей. Такие функциональные гомологи включают, например, CeresClone:1925947 (SEQ ID NO: 647), CeresAnnot:1514501 (SEQ ID NO: 649), CeresAnnot:849672 (SEQ ID NO: 651), GI:157355942 (SEQ ID NO: 652), GI:115452503 (SEQ ID NO: 653), CeresClone:1790933

(SEQ ID NO: 655), CeresAnnot:8641620 (SEQ ID NO: 657), CeresClone:281497 (SEQ ID NO: 659), GI:168013851 (SEQ ID NO: 660), CeresClone:143214 (SEQ ID NO: 662), CeresClone:1781022 (SEQ ID NO: 664), CeresClone:618639 (SEQ ID NO: 666), GI:118483001 (SEQ ID NO: 667), CeresClone:38404 (SEQ ID NO: 669), GI:3549670 (SEQ ID NO: 670), GI:37703720 (SEQ ID NO: 671), GI:152149571 (SEQ ID NO: 672), GI:125603687 (SEQ ID NO: 673), GI:108707679 (SEQ ID NO: 674), GI:157352390 (SEQ ID NO: 675), GI:159469820 (SEQ ID NO: 676), GI:145344081 (SEQ ID NO: 677) или Ceres Annot ID no. 1461228 (SEQ ID NO: 689). В некоторых случаях функциональный гомолог SEQ ID NO: 645 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 645.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 253, приведены на фиг. 6 и в перечне последовательностей. Такие функциональные гомологи включают, *например*, CeresClone:951785 (SEQ ID NO: 255), CeresAnnot:1440346 (SEQ ID NO: 257), CeresClone:1085177 (SEQ ID NO: 259) или CeresClone:157151 (SEQ ID NO: 261). В некоторых случаях функциональный гомолог с SEQ ID NO: 253 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 253.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 323, приведены на фиг. 7 и в перечне последовательностей. Такие функциональные гомологи включают, *например*, GI:157340812 (SEQ ID NO: 324), CeresAnnot:1460824 (SEQ ID NO: 326), GI:145356202 (SEQ ID NO: 327), CeresClone:477814 (SEQ ID NO: 329), CeresClone:1914387 (SEQ ID NO: 331), GI:7981380 (SEQ ID

NO: 332), CeresClone:1910072 (SEQ ID NO: 334), CeresClone:331755 (SEQ ID NO: 336), GI:124360540 (SEQ ID NO: 337), GI:157335318 (SEQ ID NO: 338), CeresAnnot:1503394 (SEQ ID NO: 340), CeresAnnot:1442707 (SEQ ID NO: 342), GI:147784500 (SEQ ID NO: 343), CeresAnnot:1514100 (SEQ ID NO: 345), CeresAnnot:850366 (SEQ ID NO: 347), CeresAnnot:543794 (SEQ ID NO: 349), CeresAnnot:1495620 (SEQ ID NO: 351), CeresClone:1653552 (SEQ ID NO: 353), GI:147767321 (SEQ ID NO: 354), CeresAnnot:1510450 (SEQ ID NO: 356), GI:110931736 (SEQ ID NO: 357), CeresClone:1916884 (SEQ ID NO: 359), CeresClone:1847251 (SEQ ID NO: 361), CeresAnnot:1457249 (SEQ ID NO: 363), CeresClone:1113584 (SEQ ID NO: 365), CeresClone:1927753 (SEQ ID NO: 367), CeresClone:857342 (SEQ ID NO: 369), CeresClone:100068619 (SEQ ID NO: 371), GI:145327247 (SEQ ID NO: 372), CeresAnnot:8461532 (SEQ ID NO: 374), CeresClone:1722230 (SEQ ID NO: 376), CeresClone:1897493 (SEQ ID NO: 378), CeresAnnot:838426 (SEQ ID NO: 380), CeresAnnot:827713 (SEQ ID NO: 382), CeresClone:1763593 (SEQ ID NO: 384), CeresClone:143475 (SEQ ID NO: 386), CeresAnnot:8456508 (SEQ ID NO: 388), CeresClone:100002959 (SEQ ID NO: 390), GI:118137433 (SEQ ID NO: 391), CeresClone:1523182 (SEQ ID NO: 393), CeresClone:1761808 (SEQ ID NO: 395), CeresClone:1069222 (SEQ ID NO: 397), CeresAnnot:8734209 (SEQ ID NO: 399), CeresAnnot:8461540 (SEQ ID NO: 401), CeresClone:1086604 (SEQ ID NO: 403), CeresClone:41695 (SEQ ID NO: 405), GI:112292440 (SEQ ID NO: 406), GI:116830269 (SEQ ID NO: 407), CeresClone:1775942 (SEQ ID NO: 409), CeresClone:1723374 (SEQ ID NO: 411), CeresAnnot:1457230 (SEQ ID NO: 413), CeresAnnot:8667653 (SEQ ID NO: 415), GI:115465643 (SEQ ID NO: 416), GI:5091605 (SEQ ID NO: 417), GI:125553458 (SEQ ID NO: 418), CeresAnnot:1510435 (SEQ ID NO: 420), GI:115438765 (SEQ ID NO: 421), GI:112292438 (SEQ ID NO: 422) или CeresAnnot:1770841 (SEQ ID NO: 424). В некоторых случаях функциональный гомолог SEQ ID NO: 323 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной

последовательности, представленной в последовательности SEQ ID NO: 323.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 595, приведены на фиг. 8 и в перечне последовательностей. Такие функциональные гомологи включают, например, CeresClone:644331 (SEQ ID NO: 597), GI:15227859 (SEQ ID NO: 598), CeresAnnot:1504349 (SEQ ID NO: 600), CeresAnnot:1265088 (SEQ ID NO: 602 (SEQ ID NO: 603), GI:125527987 (SEQ ID NO: 604), GI:14279437 (SEQ ID NO: 605), ES902065 (SEQ ID NO: 606), CeresClone:1065042 (SEQ ID NO: 608), GI:157329790 (SEQ ID NO: 609), GI:15227861 (SEQ ID NO: 610), GI:146272407 (SEQ ID NO: 611), CeresClone:95094 (SEQ ID NO: 613), CeresClone:1714893 (SEQ ID NO: 615), GI:157329890 (SEQ ID NO: 616), CeresAnnot:859635 (SEQ ID NO: 618), GI:115440397 (SEQ ID NO: 619), GI:40549303 (SEQ ID NO: 620), CeresAnnot:1457048 (SEQ ID NO: 622), GI:50401192 (SEQ ID NO: 623), CeresAnnot:1451281 (SEQ ID NO: 625), CeresAnnot:1510252 (SEQ ID NO: 627), CeresClone:1822691 (SEQ ID NO: 629), GI:197312921 (SEQ ID NO: 630), CeresAnnot:8456439 (SEQ ID NO: 632), EX096388 (SEQ ID NO: 633), GI:15028131 (SEQ ID NO: 634), CeresClone:270875 (SEQ ID NO: 636), GI:27754457 (SEQ ID NO: 637), GI:16648679 (SEQ ID NO: 638), GI:15227863 (SEQ ID NO: 639), CeresAnnot:1451282 (SEQ ID NO: 641), GI:53830670 (SEQ ID NO: 642), GI:146272405 (SEQ ID NO: 643) или CeresAnnot: 827940 (SEQ ID NO: 691). В некоторых случаях функциональный гомолог SEQ ID NO: 595 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 595.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 77, приведены на фиг. 9 и в перечне последовательностей. Такие функциональные гомологи включают, например,

CeresClone:1387948 (SEQ ID NO: 79), CeresClone:1937714 (SEQ ID NO: 81), GI:157345132 (SEQ ID NO: 82), CeresClone:464828 (SEQ ID NO: 84), CeresAnnot:1451368 (SEQ ID NO: 86), GI:37695575 (SEQ ID NO: 87), GI:116790033 (SEQ ID NO: 88), CeresClone:1346042 (SEQ ID NO: 90), CeresClone:1118610 (SEQ ID NO: 92), CeresClone:982000 (SEQ ID NO: 94), CeresClone:959670 (SEQ ID NO: 96), CeresClone:952522 (SEQ ID NO: 98), CeresClone:1914539 (SEQ ID NO: 100), CeresClone:668581 (SEQ ID NO: 102), CeresClone:1914939 (SEQ ID NO: 104), CeresClone:723694 (SEQ ID NO: 106), CeresAnnot:1456949 (SEQ ID NO: 108), CeresAnnot:1539918 (SEQ ID NO: 110), CeresAnnot:8456138 (SEQ ID NO: 112), CeresAnnot:1486506 (SEQ ID NO: 114) или GI:116786293 (SEQ ID NO: 115). В некоторых случаях функциональный гомолог SEQ ID NO: 77 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 77.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 209, приведены на фиг. 10 и в перечне последовательностей. Такие функциональные гомологи включают, например, GI:116780542 (SEQ ID NO: 210), CeresClone:1848017 (SEQ ID NO: 212), CeresAnnot:1466494 (SEQ ID NO: 214), CeresAnnot:1449022 (SEQ ID NO: 216), CeresAnnot:1482911 (SEQ ID NO: 218), CeresClone:1118987 (SEQ ID NO: 220), CeresClone:1073674 (SEQ ID NO: 222), CeresClone:1084747 (SEQ ID NO: 224), CeresClone:536345 (SEQ ID NO: 226), CeresClone:1650005 (SEQ ID NO: 228), CeresAnnot:8453882 (SEQ ID NO: 230), CeresAnnot:1373087 (SEQ ID NO: 232), CeresAnnot:8669372 (SEQ ID NO: 234), CeresClone:1048839 (SEQ ID NO: 236), CeresClone:281322 (SEQ ID NO: 238), GI:147795605 (SEQ ID NO: 239), CeresClone:2004419 (SEQ ID NO: 241), GI:125543059 (SEQ ID NO: 242), AT1G16910_LSH8 (SEQ ID NO: 243), AT1G78815_LSH7 (SEQ ID NO: 244), AT2G31160_LSH3 (SEQ ID NO: 245), AT2G42610_LSH10 (SEQ

ID NO: 246), AT3G04510_LSH2 (SEQ ID NO: 247), AT3G23290_LSH4 (SEQ ID NO: 248), AT5G28490_LSH1 (SEQ ID NO: 249), AT5G58500_LSH5 (SEQ ID NO: 250) или At1g07090_LSH6 (SEQ ID NO: 251). В некоторых случаях функциональный гомолог SEQ ID NO: 209 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 209.

Примеры аминокислотных последовательностей функциональных гомологов полипептида, представленного в последовательности SEQ ID NO: 426, приведены на фиг. 11 и в перечне последовательностей. Такие функциональные гомологи включают, например, CeresAnnot:1472338_Pb (SEQ ID NO: 428), GI:157344683_Vv (SEQ ID NO: 429), GI:87240677_Mt (SEQ ID NO: 430), GI:115448297_Os (SEQ ID NO: 431), CeresClone:1844568_Pv (SEQ ID NO: 433), CeresClone:797829_Tm (SEQ ID NO: 435), GI:168033816_Pp (SEQ ID NO: 436), GI:116788004_Ps (SEQ ID NO: 437), GI:149900503_Ha (SEQ ID NO: 438), GI:4102839_Sl (SEQ ID NO: 439), GI:31088360_Vf (SEQ ID NO: 440), CeresAnnot:8681236_Sb (SEQ ID NO: 442), CeresAnnot:8519531_Gm (SEQ ID NO: 444), CeresAnnot:8631372_Zm (SEQ ID NO: 446), GI:151426449_Hv (SEQ ID NO: 447), GI:192757675_Br (SEQ ID NO: 448), GI:2655098 (SEQ ID NO: 449), GI:194690746 (SEQ ID NO: 450), CeresClone:752925 (SEQ ID NO: 452), GI:125540898 (SEQ ID NO: 453), GI:26451333 (SEQ ID NO: 454), GI:2160144 (SEQ ID NO: 455), GI:30696666 (SEQ ID NO: 456), GI:125556922 (SEQ ID NO: 457), CeresAnnot:1529287 (SEQ ID NO: 459), CeresClone:1806748 (SEQ ID NO: 461), CeresAnnot:8755095 (SEQ ID NO: 463), GI:147827175 (SEQ ID NO: 464), CeresClone:1888865 (SEQ ID NO: 466), GI:157337163 (SEQ ID NO: 467), GI:115434472 (SEQ ID NO: 468), CeresAnnot:6252512 (SEQ ID NO: 470), CeresAnnot:1569074_Mt (SEQ ID NO: 472), CeresAnnot:1475845 (SEQ ID NO: 474), CeresAnnot:1501483 (SEQ ID NO: 476), CeresAnnot:8755079 (SEQ ID NO: 478), GI:115470147 (SEQ ID NO: 479), GI:15240905 (SEQ ID NO: 480), CeresAnnot:8755085

(SEQ ID NO: 482), GI:147853446 (SEQ ID NO: 483), GI:157346087 (SEQ ID NO: 484), CeresAnnot:1538867 (SEQ ID NO: 486), CeresAnnot:8755091 (SEQ ID NO: 488), CeresAnnot:1492702 (SEQ ID NO: 490), CeresClone:325604 (SEQ ID NO: 492), GI:108707040 (SEQ ID NO: 493), CeresAnnot:1302517_At (SEQ ID NO: 495), CeresAnnot:1355964 (SEQ ID NO: 497), CeresAnnot:8755104 (SEQ ID NO: 499), GI:147802380 (SEQ ID NO: 500), GI:510238 (SEQ ID NO: 501), GI:157341962 (SEQ ID NO: 502), GI:6635838 (SEQ ID NO: 503), GI:4455276 (SEQ ID NO: 504), CeresAnnot:8642246 (SEQ ID NO: 506), CeresAnnot:8633032 (SEQ ID NO: 508), GI:157337654 (SEQ ID NO: 509), CeresAnnot:8642241 (SEQ ID NO: 511), CeresAnnot:1520085 (SEQ ID NO: 513), CeresAnnot:1514979 (SEQ ID NO: 515), GI:147858202 (SEQ ID NO: 516), GI:125545538 (SEQ ID NO: 517), GI:115451771 (SEQ ID NO: 518), GI:125587732 (SEQ ID NO: 519), CeresAnnot:1516968 (SEQ ID NO: 521), CeresClone:350844 (SEQ ID NO: 523), CeresAnnot:8658700 (SEQ ID NO: 525), GI:157346088 (SEQ ID NO: 526), CeresClone:1926916 (SEQ ID NO: 528), GI:15226861 (SEQ ID NO: 529), CeresClone:816960 (SEQ ID NO: 531), GI:15232435 (SEQ ID NO: 532), CeresAnnot:8643789 (SEQ ID NO: 534), CeresAnnot:8631367 (SEQ ID NO: 536), GI:157339093 (SEQ ID NO: 537), CeresAnnot:8633031 (SEQ ID NO: 539), GI:125543029 (SEQ ID NO: 540), GI:115454995 (SEQ ID NO: 541), CeresAnnot:8755090 (SEQ ID NO: 543), CeresAnnot:8755097 (SEQ ID NO: 545), CeresAnnot:8755098 (SEQ ID NO: 547), CeresAnnot:8755099 (SEQ ID NO: 549) (SEQ ID NO: 550) (SEQ ID NO: 551) (SEQ ID NO: 552), CeresAnnot: 6086224 (SEQ ID NO: 693), CeresClone: 476769 (SEQ ID NO: 695) или CeresClone:15650 (SEQ ID NO: 697). В некоторых случаях функциональный гомолог SEQ ID NO: 426 имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 426.

Определение консервативных областей в модулирующем биомассу полипептиде облегчает продуцирование вариантов модулирующих

биомассу полипептидов. Варианты модулирующих биомассу полипептидов обычно имеют 10 или менее консервативных аминокислотных замен в первичной аминокислотной последовательности, например 7 или менее консервативных аминокислотных замен, 5 или менее консервативных аминокислотных замен или от 1 до 5 консервативных замен. Полезный вариант полипептида можно составить на основе одного из выравниваний, показанных на фиг. 1, фиг. 2, фиг. 3, фиг. 4, фиг. 5, фиг. 6, фиг. 7, фиг. 8, фиг. 9, фиг. 10 или фиг. 11, и/или гомологов, указанных в перечне последовательностей. Такой полипептид включает консервативные области, расположенные в порядке, показанном на фигуре, от аминоконцевой до карбоксиконцевой области. Такой полипептид может также включать ноль, одну или более одной аминокислоты в положениях, обозначенных тире. При отсутствии аминокислот в положениях, обозначенных тире, длина такого полипептида равна сумме аминокислотных остатков во всех консервативных областях. При наличии аминокислот в положении, обозначенном тире, длина такого полипептида равна сумме аминокислотных остатков во всех консервативных областях и всех тире.

В. Функциональные гомологи, определяемые с использованием программного обеспечения HMMER

В некоторых вариантах полезные модулирующие биомассу полипептиды включают полипептиды, которые совпадают с профилями скрытых марковских моделей на основе полипептидов, показанных на фиг. 1-11. Скрытая марковская модель (Hidden Markov Model, HMM) представляет собой статистическую модель консенсусной последовательности для группы функциональных гомологов. См. Durbin *et al.*, *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, UK (1998). HMM генерируется программой HMMER 2.3.2 с установленными по умолчанию параметрами и с использованием последовательностей группы функциональных гомологов в качестве входных данных. Множественное выравнивание последовательностей генерируется программой ProbCons (Do *et al.*, *Genome Res.*, 15(2):330-40 (2005)) версии 1.11 с использованием

набора параметров по умолчанию: -c, --consistency REPS из 2; -ir, --iterative-refinement REPS из 100; -pre, --pre-training REPS из 0. Программа ProbCons - это общедоступное программное обеспечение, предоставленное Стэнфордским университетом.

Для построения НММ (hmmbuild) по умолчанию используются следующие параметры: параметр по умолчанию «architecture prior» (archpri), используемый при формировании структуры MAP (белков, ассоциированных с микротрубочками), равен 0,85; порог отсека (idlelevel) по умолчанию, используемый для определения действительного номера последовательности, равен 0,62. Программа HMMER 2.3.2 была выпущена 3 октября 2003 года под универсальной общедоступной лицензией в рамках проекта по свободному распространению программного обеспечения и доступна в Интернете на различных сайтах, таких как hmm.janelia.org, hmm.wustl.edu и fr.com/hmm232/. Выходные данные модели Hmmbuild представлены в виде текстового файла.

НММ для группы функциональных гомологов можно использовать для определения вероятности того, что отбираемая последовательность модулирующего биомассу полипептида лучше соответствует конкретной НММ, чем нулевая НММ, генерированная с использованием группы последовательностей, не связанных структурно или функционально. На вероятность того, что отбираемая последовательность полипептида лучше соответствует НММ, чем нулевая НММ, указывает показатель НММ - число, генерируемое, когда отбираемая последовательность соответствует профилю НММ, который используется программой HMMER `hmmsearch`. При выполнении программы `hmmsearch` по умолчанию используются следующие параметры: значение отсека (E) по умолчанию - 10,0; показатель отсека (T) по умолчанию - минус бесконечность; количество последовательности в базе данных (Z) по умолчанию равно реальному количеству последовательностей в базе данных; значение отсека (E) по умолчанию для упорядоченного списка совпадений на каждый домен (domE) - минус бесконечность и показатель отсека для упорядоченного подоменного списка совпадений (domT) - минус бесконечность. Высокий показатель НММ указывает на более высокую вероятность того, что отбираемая

последовательность выполняет одну или несколько биохимических или физиологических функций полипептидов, используемых для генерирования НММ. Высоким показателем НММ считается показатель не менее 20, а часто даже выше. Небольшие колебания показателя НММ конкретной последовательности могут происходить вследствие некоторых факторов, таких как порядок обработки последовательностей множественными алгоритмами выравнивания последовательностей, как, например, в программе ProbCons. Однако такие колебания показателя НММ незначительны.

Модулирующие биомассу полипептиды, рассматриваемые ниже, соответствуют указанным НММ по показателю НММ, превышающему 65 (например, более 70, 80, 90, 100, 120, 140, 200, 300, 500, 1000, 1500 или 2000). В некоторых вариантах показатель НММ модулирующего биомассу полипептида, рассматриваемого ниже, составляет приблизительно 50%, 60%, 70%, 80%, 90% или 95% от показателя НММ функционального гомолога, представленного в перечне последовательностей настоящей заявки. В некоторых вариантах модулирующий биомассу полипептид, рассматриваемый ниже, соответствует указанной НММ с показателем НММ, превышающим 210, и включает домен, указывающий на модулирующий биомассу полипептид. В некоторых вариантах модулирующий биомассу полипептид, рассматриваемый ниже, соответствует указанной НММ с показателем НММ, превышающим 210, и имеет идентичность последовательности 65% или более (например, идентичность последовательности 75%, 80%, 85%, 90%, 95% или 100%) по отношению к аминокислотной последовательности, показанной на фиг. 1-11.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 130 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 1 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591 или 593.

В перечне последовательностей показаны примеры

полипептидов, у которых показатель НММ превышает 340 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 2 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320 или 321.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 530 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 3 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206 или 207.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 120 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 4 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74 или 75.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 635 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 5 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO:

645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677 или 689.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 65 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 6 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 255, 257, 259 или 261.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 100 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 7 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422 или 424.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 480 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 8 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643 или 691.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 145 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 9 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102,

104, 106, 108, 110, 112, 114 или 115.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 280 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 10 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250 или 251.

В перечне последовательностей показаны примеры полипептидов, у которых показатель НММ превышает 1000 при сопоставлении с НММ, генерированной из аминокислотных последовательностей, представленных на фиг. 11 и указанных в перечне последовательностей настоящей заявки. К таким полипептидам относятся, например, последовательности SEQ ID NO: 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 693, 695 или 697.

Г. Процент идентичности

В некоторых вариантах модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотным последовательностям, представленным в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142,

144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697. Полипептиды с таким процентом идентичности последовательности часто содержат домен, указывающий на моделирующий биомассу полипептид, и/или имеют показатель НММ, превышающий 65, как отмечалось выше. Аминокислотные последовательности моделирующих биомассу полипептидов, имеющих идентичность последовательности не менее 80% по отношению к аминокислотным последовательностям, представленным в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53,

54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613, 615, 616, 618, 619, 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637, 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657, 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675, 676, 677, 689, 691, 693, 695 или 697, представлены на фиг. 1-11 и в перечне последовательностей.

«Процент идентичности последовательности» относится к степени идентичности какой-либо заданной эталонной последовательности, например SEQ ID NO: 1, с отбираемой модулирующей биомассу последовательностью. Длина отбираемой

последовательности обычно составляет от 80 до 200 процентов длины эталонной последовательности, например 82, 85, 87, 89, 90, 93, 95, 97, 99, 100, 105, 110, 115, 120, 130, 140, 150, 160, 170, 180, 190 или 200 процентов длины эталонной последовательности. Процент идентичности какой-либо отбираемой нуклеиновой кислоты или полипептида по отношению к эталонной нуклеиновой кислоте или полипептиду можно определить следующим образом. Эталонная последовательность (например, последовательность нуклеиновой кислоты или аминокислотная последовательность) сравнивается с одной или несколькими отбираемыми последовательностями с помощью компьютерной программы ClustalW (версия 1.83, параметры по умолчанию), которая позволяет выполнять выравнивание последовательностей нуклеиновой кислоты или полипептида по всей длине (глобальное выравнивание). Chenna *et al.*, *Nucleic Acids Res.*, 31(13):3497-500 (2003).

Программа ClustalW вычисляет лучшее совпадение между эталонной и одной или несколькими отбираемыми последовательностями и сопоставляет их таким образом, чтобы можно было определить идентичность, сходства и различия. Для максимального выравнивания последовательностей делеции в одном или нескольких остатках можно вставить в эталонную последовательность, отбираемую последовательность или в обе последовательности. Для быстрого парного выравнивания последовательностей нуклеиновой кислоты используются следующие параметры по умолчанию: разрядность: 2; размер окна: 4; метод количественной оценки: процентное соотношение; количество верхних диагоналей: 4; штраф за делецию: 5. Для множественного выравнивания последовательностей нуклеиновых кислот используются следующие параметры: штраф за открытие делеции: 10,0; штраф за продолжение делеции: 5,0; вес перехода: да. Для быстрого парного выравнивания белковых последовательностей используются следующие параметры: разрядность: 1; размер окна: 5; метод количественной оценки: процентное соотношение; количество верхних диагоналей: 5; штраф за делецию: 3. Для множественного выравнивания белковых последовательностей используются следующие параметры: матрица

веса: *blosum*; штраф за открытие делеции: 10,0; штраф за продолжение делеции: 0,05; гидрофильные делеции: *on*; гидрофильные остатки: *Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg* и *Lys*; штрафы за остатки в области делеции: *on*. Выходные данные программы *ClustalW* представляются в виде выравнивания последовательности, которое отражает связи между последовательностями. Программу *ClustalW* можно запустить в Интернете, например, на поисковом узле веб-сайта Медицинского колледжа Вэйлора (searchlauncher.bcm.tmc.edu/multi-align/multi-align.html) и на веб-сайте Европейского института биоинформатики (ebi.ac.uk/clustalw).

Для определения процента идентичности отбираемой нуклеиновой кислоты или аминокислотной последовательности по отношению к эталонной последовательности, последовательности выравниваются с помощью программы *ClustalW*, количество идентичных совпадений в выравнивании делится на длину эталонной последовательности и результат умножается на 100. Следует отметить, что значение процента идентичности можно округлить до ближайшей десятой доли. Например, 78,11, 78,12, 78,13 и 78,14 округляются до 78,1, в то время как 78,15, 78,16, 78,7, 78,18 и 78,19 округляются до 78,2.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например, идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 554. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 554, приведены на фиг. 1 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной

последовательности, представленной в последовательности SEQ ID NO: 263. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 263, приведены на фиг. 2 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 117. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 117, приведены на фиг. 3 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 1. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 1, приведены на фиг. 4 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 645. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 645, приведены на фиг. 5 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет

аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 253. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 253, приведены на фиг. 6 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 323. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 323, приведены на фиг. 7 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 595. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 595, приведены на фиг. 8 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, *например* идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 77. Аминокислотные последовательности полипептидов, имеющие

идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 77, приведены на фиг. 9 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 209. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 09, приведены на фиг. 10 и в перечне последовательностей.

В некоторых случаях модулирующий биомассу полипептид имеет аминокислотную последовательность с идентичностью последовательности не менее 45%, например идентичность последовательности 50%, 52%, 56%, 59%, 61%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к аминокислотной последовательности, представленной в последовательности SEQ ID NO: 426. Аминокислотные последовательности полипептидов, имеющие идентичность последовательности более 45% по отношению к полипептиду, представленному в последовательности SEQ ID NO: 426, приведены на фиг. 11 и в перечне последовательностей.

Д. Прочие последовательности

Следует отметить, что модулирующий биомассу полипептид может включать дополнительные аминокислоты, которые не участвуют в модуляции биомассы, и поэтому такой полипептид может быть длиннее, чем обычно. Например, модулирующий биомассу полипептид может включать метку очистки, транзитный пептид хлоропласта, митохондриальный транзитный пептид, пептид аминоконцевого или лидерную последовательность, добавленную к аминоконцевой или карбоксиконцевой области. В некоторых вариантах модулирующий биомассу полипептид включает аминокислотную последовательность, которая выполняет функцию репортера, например зеленый флуоресцентный белок или желтый флуоресцентный белок.

III. Нуклеиновые кислоты

Нуклеиновые кислоты, описанные в настоящем документе, включают нуклеиновые кислоты, которые эффективны для модуляции уровней биомассы при транскрибировании в растении или растительной клетке. Такие нуклеиновые кислоты в том числе включают те нуклеиновые кислоты, которые кодируют модулирующий биомассу полипептид, и те, которые могут быть использованы для ингибирования экспрессии модулирующего биомассу полипептида нуклеиновой кислотой.

А. Нуклеиновые кислоты, которые кодируют модулирующие биомассу полипептиды

В настоящем документе описаны нуклеиновые кислоты, которые кодируют модулирующие биомассу полипептиды. Примерами таких нуклеиновых кислот являются последовательности SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696, как более подробно описано ниже. Нуклеиновая кислота также может быть фрагментом, который составляет не менее 40% (например, не менее 45, 50, 55, 60, 65,

70, 75, 80, 85, 90, 95 или 99%) длины полноразмерной нуклеиновой кислоты, приведенной в последовательностях SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28, 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76, 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143, 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181, 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256, 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288, 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330, 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362, 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389, 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423, 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469, 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505, 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542, 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572, 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601, 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646, 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 553. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 553. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 553.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в

последовательности SEQ ID NO: 262. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 262. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 262.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 116. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 116. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 116.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 678. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 678. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 678.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в

последовательности SEQ ID NO: 644. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 644. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 644.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 252. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 252. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 252.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 322. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 322. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 322.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в

последовательности SEQ ID NO: 594. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 594. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 594.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 76. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 76. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 76.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 208. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 208. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 208.

Модулирующая биомассу нуклеиновая кислота может содержать нуклеотидную последовательность, представленную в

последовательности SEQ ID NO: 425. Или же модулирующая биомассу нуклеиновая кислота может быть вариантом нуклеиновой кислоты, имеющей нуклеотидную последовательность, представленную в последовательности SEQ ID NO: 425. Например, модулирующая биомассу нуклеиновая кислота может иметь нуклеотидную последовательность с идентичностью последовательности не менее 80%, например идентичность последовательности 81%, 85%, 90%, 95%, 97%, 98% или 99% по отношению к нуклеотидной последовательности, представленной в последовательности SEQ ID NO: 425.

Изолированные молекулы нуклеиновой кислоты могут быть получены стандартными способами. Например, для получения изолированной нуклеиновой кислоты, которая содержит нуклеотидную последовательность, описанную в настоящем документе, можно использовать метод полимеразной цепной реакции (ПЦР). ПЦР можно использовать для амплификации специфических последовательностей как ДНК, так и РНК, в том числе последовательностей всей геномной ДНК или всей клеточной РНК. Различные методы ПЦР описаны, например, в праймере для ПЦР: A Laboratory Manual, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995. Как правило, для конструирования олигонуклеотидных праймеров, которые в последовательности идентичны или подобны противоположащим цепям амплифицируемой матрицы, используется информация о последовательности от концов области, представляющей интерес, или из участков, находящихся за ее пределами. Также имеются различные стратегии с применением ПЦР, посредством которых сайт-специфичные модификации нуклеотидных последовательностей могут быть интродуцированы в матричную нуклеиновую кислоту. Изолированные нуклеиновые кислоты также могут быть синтезированы химическим путем или в виде отдельной молекулы нуклеиновой кислоты (например, фосфорамидитным методом с использованием автоматического синтеза ДНК в направлении от 3' к 5'), или в виде серии олигонуклеотидов. Например, может быть синтезирована одна или несколько пар длинных олигонуклеотидов (например, >100 нуклеотидов), которые содержат требуемую последовательность, при этом каждая пара содержит короткий

комплементарный участок (например, приблизительно 15 нуклеотидов), в котором при отжиге пары олигонуклеотидов формируется дуплекс. Для вытягивания олигонуклеотидов используется ДНК-полимераза, и в результате на каждой паре олигонуклеотидов формируется отдельная молекула двухцепочечной нуклеиновой кислоты, которая затем может быть лигирована в вектор. Изолированные нуклеиновые кислоты согласно настоящему изобретению можно получить путем мутагенеза, например, природной ДНК.

Б. Использование нуклеиновых кислот для модуляции экспрессии полипептидов

i. Экспрессия модулирующего биомассу полипептида

Нуклеиновая кислота, кодирующая один из модулирующих биомассу полипептидов, которые описаны в настоящем документе, может быть использована для экспрессии полипептида в изучаемом виде растения, как правило, путем трансформирования растительной клетки с помощью нуклеиновой кислоты, имеющей кодирующую последовательность полипептида, функционально связанную в смысловой ориентации с одной или более регуляторными областями. Следует иметь в виду, что из-за вырожденности генетического кода ряд нуклеиновых кислот может кодировать определенный модулирующий биомассу полипептид; то есть для многих аминокислот существует более чем один нуклеотидный триплет, который служит кодоном аминокислоты. Таким образом, для получения оптимальной экспрессии в определенном виде растения можно модифицировать кодоны в кодирующей последовательности заданного модулирующего биомассу полипептида, используя соответствующие таблицы статистического отклонения от равномерности в использовании кодонов для данного вида.

В некоторых случаях экспрессия модулирующего биомассу полипептида ингибирует одну или несколько функций эндогенного полипептида. Например, нуклеиновую кислоту, которая кодирует доминантно-негативный полипептид, можно использовать для ингибирования функции белка. Доминантно-негативный полипептид обычно мутирован или усечен по сравнению с эндогенным полипептидом дикого типа, и его присутствие в клетке ингибирует

одну или несколько функций полипептида дикого типа в данной клетке, то есть доминантно-негативный полипептид является генетически доминантным и передает потерю функции. Механизм, посредством которого доминантно-негативный полипептид передает такой фенотип, может быть различным, однако часто может включать белок-белковое взаимодействие или взаимодействие белок-ДНК. Например, доминантно-негативный полипептид может быть ферментом, усеченным по сравнению с нативным ферментом дикого типа, так что усеченный полипептид удерживает домены, участвующие в связывании первого белка, однако испытывает нехватку доменов, задействованных в связывании второго белка. Поэтому усеченный полипептид не способен правильно модулировать деятельность второго белка. См., например, патентную заявку US 2007/0056058. Другой пример: точечная мутация, при которой происходит замена неконсервативных аминокислот в каталитическом домене, может привести к появлению доминантно-негативного полипептида. См., например, патентную заявку US 2005/032221. Еще один пример: доминантно-негативный полипептид может являться фактором транскрипции, усеченным по сравнению с нативным фактором транскрипции дикого типа, так что усеченный полипептид удерживает ДНК-связывающий домен (ы), однако испытывает нехватку домена (ов) активации. Такой усеченный полипептид может ингибировать связывание ДНК фактором транскрипции дикого типа, ингибируя тем самым активацию транскрипции.

ii. Ингибирование экспрессии модулирующего биомассу полипептида

Полинуклеотиды и рекомбинантные структуры, описанные в настоящем документе, могут быть использованы для ингибирования экспрессии модулирующего биомассу полипептида в изучаемом виде растения. См., например, Matzke and Birchler, Nature Reviews Genetics 6:24-35 (2005); Akashi et al., Nature Reviews Mol. Cell Biology 6:413-422 (2005); Mittal, Nature Reviews Genetics 5:355-365 (2004) и Nature Reviews RNA interference collection, Oct. 2005 в Интернете на веб-сайте nature.com/reviews/focus/mai. Известен ряд методов ингибирования экспрессии генов в растениях с помощью нуклеиновых кислот, в том числе расщепление

антисмысловой РНК и рибозим-направляемой РНК, посттранскрипционное подавление активности гена (PTGS), например РНК-интерференция (РНКи) и транскрипционное подавление активности гена (TGS). Подходящие полинуклеотиды включают полноразмерные нуклеиновые кислоты, которые кодируют модулирующие биомассу полипептиды, или фрагменты таких полноразмерных нуклеиновых кислот. В некоторых вариантах может быть использован комплемент полноразмерной нуклеиновой кислоты или ее фрагмент. Обычно фрагмент включает не менее 10 нуклеотидов, например не менее 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 35, 40, 50, 80, 100, 200, 500 нуклеотидов или более. В целом, для компенсации более короткой последовательности может быть использована гомология более высокого уровня.

Один из общеизвестных методов – это антисмысловая технология. В данном методе нуклеиновая кислота репрессируемого гена клонируется и функционально связывается с регуляторной областью и терминирующей транскрипцию последовательностью таким образом, что происходит транскрибирование антисмысловой цепи РНК. Затем происходит изменение рекомбинантной структуры в растениях, как описано в настоящем документе, и продуцируется антисмысловая цепь РНК. Нуклеиновая кислота не должна быть всей последовательностью репрессируемого гена, однако, как правило, она будет существенным дополнением по меньшей мере к участку смысловой цепи репрессируемого гена.

В другом методе нуклеиновая кислота может быть транскрибирована в рибозим или каталитическую РНК, которая влияет на экспрессию мРНК. См. патент США № 6,423,885. Рибозимы могут быть сконструированы в виде особой пары фактически с любой РНК-мишенью, и они могут расщепить фосфодиэфирный остов в конкретном месте, функционально инактивируя таким образом РНК-мишень. Гетерологичные нуклеиновые кислоты могут кодировать рибозимы, сконструированные для расщепления конкретных транскриптов мРНК, предотвращая таким образом экспрессию полипептида. Для разрушения конкретных мРНК применяются рибозимы типа «головка молотка», хотя могут быть использованы и различные

рибозимы, которые расщепляют мРНК в сайт-специфичных последовательностях распознавания. Рибозимы типа «головка молотка» расщепляют мРНК в местах, определяемых фланкирующими областями, которые формируют комплементарные пары оснований с мРНК-мишенью. Единственное требование состоит в том, что РНК-мишень должна содержать нуклеотидную последовательность 5'-UG-3'.

Конструирование и продуцирование рибозимов типа «головка молотка» в данной отрасли знаний известны. См., например, патент США № 5,254,678 и документ WO 02/46449, а также приведенные в них материалы, использованные при экспертизе заявки. Для повышения эффективности расщепления *in vivo* в стабильную РНК, такую как транспортная РНК (тРНК), можно внедрить последовательности рибозимов типа «головка молотка». Perriman *et al.*, *Proc. Natl. Acad. Sci. USA*, 92(13):6175-6179 (1995); de Feyter and Gaudron, *Methods in Molecular Biology*, Vol. 74, Chapter 43, "Expressing Ribozymes in Plants", Edited by Turner, P.C., Humana Press Inc., Totowa, NJ. Могут быть использованы описанные эндорибонуклеазы РНК, например, встречающиеся в природе в *Tetrahymena thermophila*. См., например, патенты США № 4,987,071 и 6,423,885.

Для ингибирования экспрессии гена может быть также использовано, например, посттранскрипционное подавление активности гена (PTGS). Например, можно подготовить конструкцию, включающую последовательность, которая транскрибируется в РНК и может вызвать самоотжиг, например, с образованием двухцепочечной РНК, имеющей структуру типа «стебель-петля». В некоторых вариантах одна цепь стволового участка двухцепочечной РНК включает последовательность, подобную или идентичную смысловой кодирующей последовательности (или ее фрагменту) модулирующего биомассу полипептида длиной приблизительно от 10 до 2500 нуклеотидов. Длина последовательности, подобной или идентичной смысловой кодирующей последовательности, может быть от 10 до 500 нуклеотидов, от 15 до 300 нуклеотидов, от 20 до 100 нуклеотидов или от 25 до 100 нуклеотидов. Другая цепь стволового участка двухцепочечной РНК включает последовательность, подобную или

идентичную антисмысловой цепи (или ее фрагменту) кодирующей последовательности модулирующего биомассу полипептида, и может иметь длину короче, такую же или больше, чем длина соответствующей смысловой последовательности. В некоторых случаях одна цепь стволового участка двухцепочечной РНК включает последовательность, подобную или идентичную нетранслируемой области 3' или 5' (или ее фрагменту) мРНК, которая кодирует модулирующий биомассу полипептид, а другая цепь стволового участка двухцепочечной РНК включает последовательность, подобную или идентичную последовательности, которая является, соответственно, дополнением нетранслируемой области 3' или 5' (или ее фрагмента) мРНК, которая кодирует модулирующий биомассу полипептид. В других вариантах одна цепь стволового участка двухцепочечной РНК включает последовательность, подобную или идентичную последовательности интрона (или его фрагменту) в пре-мРНК, которая кодирует модулирующий биомассу полипептид, а другая цепь стволового участка включает последовательность, подобную или идентичную последовательности, которая является дополнением к последовательности интрона (или его фрагменту) в пре-мРНК.

Петлевой участок двухцепочечной РНК может включать от 3 до 5000 нуклеотидов, например от 3 до 25 нуклеотидов, от 15 до 1000 нуклеотидов, от 20 до 500 нуклеотидов или от 25 до 200 нуклеотидов. Петлевой участок РНК может включать интрон или его фрагмент. Двухцепочечная РНК может иметь ноль, один, два, три, четыре, пять, шесть, семь, восемь, девять, десять или более структур с петлевыми участками.

Структура, включающая последовательность, которая функционально связана с регуляторной областью и терминирующей транскрипцию последовательностью и которая транскрибируется в РНК, которая может формировать двухцепочечную РНК, трансформируется в растениях, как описано в настоящем документе. Специалистам в данной отрасли знаний известны методы использования РНКи для ингибирования экспрессии гена. См., например, патенты США № 5,034,323, 6,326,527, 6,452,067, 6,573,099, 6,753,139 и 6,777,588. См. также документы WO

97/01952, WO 98/53083, WO 99/32619, WO 98/36083 и публикации патентов США 20030175965, 20030175783, 20040214330 и 20030180945.

Структуры, которые содержат регуляторные области, функционально связанные с молекулами нуклеиновой кислоты в смысловой ориентации, могут быть также использованы для ингибирования экспрессии гена. Продукт транскрипции может быть подобен или идентичен смысловой кодирующей последовательности (или ее фрагменту) модулирующего биомассу полипептида. Продукт транскрипции может быть также неполиаденилированным, в нем может отсутствовать 5'кэп-структура, или он может содержать несплайсируемый интрон. Методы ингибирования экспрессии гена с использованием полноразмерной кДНК, а также частичной последовательности кДНК, известны в данной отрасли знаний. См., например, патент США № 5,231,020.

В некоторых вариантах структура с нуклеиновой кислотой, имеющей как минимум одну цепь, которая является матрицей и для смысловой, и для несмысловой последовательностей, комплементарных друг другу, используется для ингибирования экспрессии гена. Смысловая и антисмысловая последовательности могут быть частью большей молекулы нуклеиновой кислоты или могут быть частью отдельных молекул нуклеиновой кислоты, последовательности которых не являются комплементарными. Смысловая или антисмысловая последовательность может быть последовательностью, которая идентична (или комплементарна) последовательности мРНК, нетранслируемой областью 3' или 5' мРНК, или интрона в пре-мРНК, которая кодирует модулирующий биомассу полипептид, или фрагменту таких последовательностей. В некоторых вариантах смысловая или несмысловая последовательность идентична или комплементарна последовательности регуляторной области, которая стимулирует транскрипцию гена, кодирующего модулирующий биомассу полипептид. В каждом случае смысловая последовательность является последовательностью, комплементарной антисмысловой последовательности.

Длина смысловой и антисмысловой последовательностей может быть больше примерно 10 нуклеотидов (например, 12, 13, 14, 15,

16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более нуклеотидов). Например, антисмысловая последовательность может быть длиной 21 или 22 нуклеотида. Обычно смысловая и антисмысловая последовательности имеют длину примерно от 15 до 30 нуклеотидов, например, примерно от 18 до 28 нуклеотидов или примерно от 21 до 25 нуклеотидов.

В некоторых вариантах антисмысловая последовательность является последовательностью, комплементарной последовательности мРНК (или ее фрагменту), которая кодирует модулирующий биомассу полипептид, описанный в настоящем документе. Смысловая последовательность, комплементарная антисмысловой последовательности, может являться последовательностью, присутствующей в мРНК кодирующей биомассу полипептида. Как правило, смысловая и антисмысловая последовательности сконструированы таким образом, чтобы соответствовать 15-30 нуклеотидной последовательности мРНК-мишени для того, чтобы понизить уровень мРНК-мишени.

В некоторых вариантах для ингибирования экспрессии гена может быть использована структура с нуклеиновой кислотой, имеющей как минимум одну цепь, которая является матрицей для нескольких смысловых последовательностей (например, для 2, 3, 4, 5, 6, 7, 8, 9, 10 или более смысловых последовательностей). Аналогичным образом, для ингибирования экспрессии гена может быть использована структура с нуклеиновой кислотой, имеющей как минимум одну цепь, которая является матрицей для нескольких антисмысловых последовательностей (например, для 2, 3, 4, 5, 6, 7, 8, 9, 10 или более антисмысловых последовательностей). Например, конструкция может содержать нуклеиновую кислоту, имеющую как минимум одну цепь, которая является матрицей и для двух смысловых последовательностей, и для двух антисмысловых последовательностей. Множественные смысловые последовательности могут быть идентичными или различными, и множественные антисмысловые последовательности могут быть идентичными или различными. Например, конструкция может иметь нуклеиновую кислоту, имеющую одну цепь, которая является матрицей для двух идентичных смысловых последовательностей и двух идентичных

антисмысловых последовательностей, которые комплементарны двум идентичным смысловым последовательностям. Или же изолированная нуклеиновая кислота может иметь одну цепь, которая является матрицей для (1) двух идентичных смысловых последовательностей длиной 20 нуклеотидов, (2) одной антисмысловой последовательности, комплементарной двум идентичным смысловым последовательностям длиной 20 нуклеотидов, (3) смысловой последовательности длиной 30 нуклеотидов и (4) трех идентичных антисмысловых последовательностей, комплементарных смысловой последовательности длиной 30 нуклеотидов. Конструкции, приведенные в настоящем документе, могут быть составлены таким образом, чтобы они были пригодны для смысловых или антисмысловых последовательностей. Например, за двумя идентичными смысловыми последовательностями могут следовать две идентичные антисмысловые последовательности, или их можно расположить между двумя идентичными антисмысловыми последовательностями.

Нуклеиновая кислота, имеющая как минимум одну цепь, которая является матрицей для одной или нескольких смысловых и/или антисмысловых последовательностей, может быть функционально связана с регуляторной областью для стимуляции транскрипции молекулы РНК, содержащей смысловую и/или антисмысловую последовательность (и). Кроме того, такая нуклеиновая кислота может быть функционально связана с терминирующей транскрипцию последовательностью, такой как терминатор гена нопалинсентазы (*nos*). В некоторых случаях две регуляторные области могут направлять транскрипцию двух транскриптов: одну из верхней цепи и одну из нижней цепи. См. например, Yan *et al.*, *Plant Physiol.*, 141:1508-1518 (2006). Две регуляторные области могут быть одинаковыми или различными. Два транскрипта могут образовывать молекулы двухцепочечной РНК, которые индуцируют деградацию РНК-мишени. В некоторых случаях нуклеиновую кислоту можно расположить в пределах Т-ДНК или выделенной из растения перенесенной ДНК (пахитенной ДНК) таким образом, чтобы последовательности левой и правой границы Т-ДНК или последовательности левой и правой пахитенной ДНК прилегали к нуклеиновой кислоте или находились на одной из сторон

нуклеиновой кислоты. См. патентную заявку US 2006/0265788. Последовательность нуклеиновой кислоты между двумя регуляторными областями может быть длиной примерно от 15 до 300 нуклеотидов. В некоторых вариантах последовательность нуклеиновой кислоты между двумя регуляторными областями имеет длину примерно от 15 до 200 нуклеотидов, примерно от 15 до 100 нуклеотидов, примерно от 15 до 50 нуклеотидов, примерно от 18 до 50 нуклеотидов, примерно от 18 до 40 нуклеотидов, примерно от 18 до 30 нуклеотидов или примерно от 18 до 25 нуклеотидов.

В некоторых методах ингибирования экспрессии гена у растений, основанных на использовании нуклеиновой кислоты, подходящей нуклеиновой кислотой может быть аналог нуклеиновой кислоты. Аналоги нуклеиновой кислоты можно модифицировать на базовом компоненте, сахарном компоненте или фосфатном остове, например, для улучшения стабильности, гибридизации или растворимости нуклеиновой кислоты. Модификации на базовом компоненте включают превращения дезоксиуридина в дезокситимидин, а также 5-метил-2'-дезоксцитидин и 5-бромо-2'-дезоксцитидин в дезоксцитидин. Модификации на сахарном компоненте включают превращении 2' гидроксила сахара рибозы в форму сахаров 2'-O-метил или 2'-O-аллил. Можно модифицировать фосфатный остов дезоксирибозы для продуцирования морфолинонуклеиновых кислот, у которых каждый базовый компонент связан с шестичленным морфолиновым кольцом, или пептидных нуклеиновых кислот, у которых остов дезоксифосфата заменен на псевдопептидный остов, а четыре основания сохраняются. См., например, Summerton and Weller, *Antisense Nucleic Acid Drug Dev.*, 7:187-195 (1997); Nyrup et al., *Bioorgan. Med. Chem.*, 4:5-23 (1996). Кроме того, остов дезоксифосфата можно заменить, например, на фосфоротриат или остов фосфоритриата, фосфоамидит или остов алкилфосфотриэфира.

В. Конструкции/векторы

Рекомбинантные конструкции, описанные в настоящем документе, могут быть использованы для трансформации растений или растительных клеток с целью модулирования уровней биомассы. Как описано в настоящем документе, рекомбинантная конструкция

нуклеиновой кислоты может включать нуклеиновую кислоту, которая кодирует модулирующий биомассу полипептид, функционально связанный с регуляторной областью, пригодной для того, чтобы экспрессировать модулирующий биомассу полипептид в растении или клетке. Таким образом, нуклеиновая кислота может включать кодирующую последовательность, которая кодирует модулирующие биомассу полипептиды, как описано в последовательностях SEQ ID NO: 1, 2, 4, 6, 8, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 22, 24, 25, 27, 29, 30, 32, 33, 34, 36, 37, 38, 39, 40, 41, 43, 45, 47, 49, 50, 51, 53, 54, 56, 58, 59, 61, 63, 64, 66, 68, 70, 71, 72, 74, 75, 77, 79, 81, 82, 84, 86, 87, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 115, 117, 118, 120, 121, 122, 123, 125, 127, 129, 131, 132, 133, 135, 137, 139, 141, 142, 144, 145, 146, 147, 149, 151, 152, 153, 154, 155, 156, 158, 160, 162, 163, 164, 166, 168, 169, 171, 173, 174, 176, 178, 180, 182, 184, 185, 186, 188, 189, 190, 191, 193, 194, 195, 196, 198, 200, 202, 203, 204, 206, 207, 209, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 239, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 253, 255, 257, 259, 261, 263, 264, 266, 268, 269, 271, 273, 275, 276, 278, 279, 281, 282, 283, 285, 287, 289, 291, 292, 294, 295, 296, 297, 298, 299, 300, 302, 304, 305, 306, 308, 310, 311, 312, 314, 315, 317, 319, 320, 321, 323, 324, 326, 327, 329, 331, 332, 334, 336, 337, 338, 340, 342, 343, 345, 347, 349, 351, 353, 354, 356, 357, 359, 361, 363, 365, 367, 369, 371, 372, 374, 376, 378, 380, 382, 384, 386, 388, 390, 391, 393, 395, 397, 399, 401, 403, 405, 406, 407, 409, 411, 413, 415, 416, 417, 418, 420, 421, 422, 424, 426, 428, 429, 430, 431, 433, 435, 436, 437, 438, 439, 440, 442, 444, 446, 447, 448, 449, 450, 452, 453, 454, 455, 456, 457, 459, 461, 463, 464, 466, 467, 468, 470, 472, 474, 476, 478, 479, 480, 482, 483, 484, 486, 488, 490, 492, 493, 495, 497, 499, 500, 501, 502, 503, 504, 506, 508, 509, 511, 513, 515, 516, 517, 518, 519, 521, 523, 525, 526, 528, 529, 531, 532, 534, 536, 537, 539, 540, 541, 543, 545, 547, 549, 550, 551, 552, 554, 556, 558, 560, 562, 563, 565, 567, 569, 571, 573, 574, 575, 577, 579, 581, 583, 585, 587, 589, 591, 593, 595, 597, 598, 600, 602, 603, 604, 605, 606, 608, 609, 610, 611, 613,

615, 616, 618, 619,
 620, 622, 623, 625, 627, 629, 630, 632, 633, 634, 636, 637,
 638, 639, 641, 642, 643, 645, 647, 649, 651, 652, 653, 655, 657,
 659, 660, 662, 664, 666, 667, 669, 670, 671, 672, 673, 674, 675,
 676, 677, 689, 691, 693, 695 или 697. Примеры нуклеиновых
 кислот, кодирующих модулирующие биомассу полипептиды, приведены
 в последовательностях SEQ ID NO: 3, 5, 7, 9, 19, 21, 23, 26, 28,
 31, 35, 42, 44, 46, 48, 52, 55, 57, 60, 62, 65, 67, 69, 73, 76,
 78, 80, 83, 85, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109,
 111, 113, 116, 119, 124, 126, 128, 130, 134, 136, 138, 140, 143,
 148, 150, 157, 159, 161, 165, 167, 170, 172, 175, 177, 179, 181,
 183, 187, 192, 197, 199, 201, 205, 208, 211, 213, 215, 217, 219,
 221, 223, 225, 227, 229, 231, 233, 235, 237, 240, 252, 254, 256,
 258, 260, 262, 265, 267, 270, 272, 274, 277, 280, 284, 286, 288,
 290, 293, 301, 303, 307, 309, 313, 316, 318, 322, 325, 328, 330,
 333, 335, 339, 341, 344, 346, 348, 350, 352, 355, 358, 360, 362,
 364, 366, 368, 370, 373, 375, 377, 379, 381, 383, 385, 387, 389,
 392, 394, 396, 398, 400, 402, 404, 408, 410, 412, 414, 419, 423,
 425, 427, 432, 434, 441, 443, 445, 451, 458, 460, 462, 465, 469,
 471, 473, 475, 477, 481, 485, 487, 489, 491, 494, 496, 498, 505,
 507, 510, 512, 514, 520, 522, 524, 527, 530, 533, 535, 538, 542,
 544, 546, 548, 553, 555, 557, 559, 561, 564, 566, 568, 570, 572,
 576, 578, 580, 582, 584, 586, 588, 590, 592, 594, 596, 599, 601,
 607, 612, 614, 617, 621, 624, 626, 628, 631, 635, 640, 644, 646,
 648, 650, 654, 656, 658, 661, 663, 665, 668, 678, 679, 680, 681,
 682, 683, 684, 685, 686, 687, 688, 690, 692, 694 или 696 или в
 перечне последовательностей. Модулирующий биомассу полипептид,
 закодированный рекомбинантной нуклеиновой кислотой, может быть
 нативным модулирующим биомассу полипептидом или может быть
 гетерологичным по отношению к клетке. В некоторых случаях
 рекомбинантная конструкция содержит нуклеиновую кислоту, которая
 ингибирует экспрессию модулирующего биомассу полипептида,
 функционально связанного с регуляторной областью. Примеры
 подходящих регуляторных областей описаны в разделе «Регуляторные
 области».

Также представлены векторы, содержащие конструкции

рекомбинантных нуклеиновых кислот, аналогичных тем, которые описаны в настоящем документе. Подходящими остовами векторов являются, например, те, которые обычно используются в данной отрасли, такие как плазмиды, вирусы, искусственные хромосомы, искусственные бактериальные хромосомы, искусственные дрожжевые хромосомы или искусственные хромосомы P1. Подходящими экспрессионными векторами являются в том числе плазмиды и вирусные векторы, полученные, например, из бактериофагов, бакуловирусов и ретровирусов. Многочисленные векторы и экспрессирующие системы доступны для приобретения в таких корпорациях, как Novagen® (Мэдисон, штат Висконсин), Clontech® (Пало-Альто, штат Калифорния), Stratagene® (Ла-Хойя, штат Калифорния) и Invitrogen/Life Technologies® (Карлсбад, штат Калифорния).

Векторы, приводимые в настоящем документе, также могут включать, например, области начала репликации, области прикрепления к скелету и/или маркеры.

Маркерный ген может передавать растительной клетке селективируемый фенотип. Например, маркер может передавать устойчивость к биоцидам, такую как устойчивость к антибиотикам (например, канамицину, G418, блеомицину или гигромицину), или к гербицидам (например, глифосату, хлорсульфурону или фосфинотрицину). Кроме того, экспрессионный вектор может включать маркерную последовательность, предназначенную для облегчения манипуляции с экспрессируемыми полипептидами или их обнаружения (например, очистки или локализации). Маркерные последовательности, такие как люцифераза, β -глюкуронидаза (GUS), зеленый флуоресцентный белок (GFP), глутатион S-трансфераза (GST), полигистидин, с-мус, гемагглютинин или маркер Flag™ (Kodak, Нью-Хейвен, штат Коннектикут), обычно экспрессируются в форме слияния с кодированным полипептидом. Такие маркеры можно вставлять в любом месте полипептида, в том числе в аминоконцевой или карбоксиконцевой области.

Г. Регуляторные области

Выбор регуляторных областей для включения в рекомбинантную конструкцию зависит от нескольких факторов, в том числе от

эффективности, селективности, способности к индуцированию, желательного уровня экспрессии и преимущественно клеточной или тканевой экспрессии. Для специалиста в данной области обычным делом является модулирование экспрессии кодирующей последовательности путем соответствующего подбора и расположения регуляторных областей в кодирующей последовательности. Транскрипцию нуклеиновой кислоты можно модулировать аналогичным образом.

Некоторые подходящие регуляторные области инициируют транскрипцию только или преимущественно в клетках определенных типов. Методы идентификации и определения характеристик регуляторных областей в геномной ДНК растения известны и описаны в том числе в следующих ссылочных материалах: Jordano *et al.*, *Plant Cell*, 1:855-866 (1989); Bustos *et al.*, *Plant Cell*, 1:839-854 (1989); Green *et al.*, *EMBO J.*, 7:4035-4044 (1988); Meier *et al.*, *Plant Cell*, 3:309-316 (1991); и Zhang *et al.*, *Plant Physiology*, 110:1069-1079 (1996).

Примеры различных классов регуляторных областей описаны ниже. Некоторые регуляторные области, указанные ниже, а также дополнительные регуляторные области более подробно описаны в заявках на патент США с серийными номерами 60/505,689; 60/518,075; 60/544,771; 60/558,869; 60/583,691; 60/619,181; 60/637,140; 60/757,544; 60/776,307; 10/957,569; 11/058,689; 11/172,703; 11/208,308; 11/274,890; 60/583,609; 60/612,891; 11/097,589; 11/233,726; 11/408,791; 11/414,142; 10/950,321; 11/360,017; PCT/US05/011105; PCT/US05/23639; PCT/US05/034308; PCT/US05/034343; а также PCT/US06/038236; PCT/US06/040572 и PCT/US07/62762.

Например, последовательности регуляторных областей p326, YP0144, YP0190, p13879, YP0050, p32449, 21876, YP0158, YP0214, YP0380, PT0848, PT0633, YP0128, YP0275, PT0660, PT0683, PT0758, PT0613, PT0672, PT0688, PT0837, YP0092, PT0676, PT0708, YP0396, YP0007, YP0111, YP0103, YP0028, YP0121, YP0008, YP0039, YP0115, YP0119, YP0120, YP0374, YP0101, YP0102, YP0110, YP0117, YP0137, YP0285, YP0212, YP0097, YP0107, YP0088, YP0143, YP0156, PT0650, PT0695, PT0723, PT0838, PT0879, PT0740, PT0535, PT0668, PT0886,

PT0585, YP0381, YP0337, PT0710, YP0356, YP0385, YP0384, YP0286, YP0377, PD1367, PT0863, PT0829, PT0665, PT0678, YP0086, YP0188, YP0263, PT0743 и YP0096 приведены в перечне последовательностей PCT/US06/040572; последовательность регуляторной области PT0625 приведена в перечне последовательностей PCT/US05/034343; последовательности регуляторных областей PT0623, YP0388, YP0087, YP0093, YP0108, YP0022 и YP0080 приведены в перечне последовательностей в заявке на патент США с серийным номером 11/172,703; последовательность регуляторной области PR0924 приведена в перечне последовательностей PCT/US07/62762; и последовательности регуляторных областей p530c10, pOsFIE2-2, pOsMEA, pOsYp102 и pOsYp285 приведены в перечне последовательностей PCT/US06/038236.

Следует иметь в виду, что регуляторная область может удовлетворять критериям одной классификации, исходя из ее действия в растениях одного вида, и одновременно удовлетворять критериям другой классификации, исходя из ее действия в растениях другого вида.

i. Экспрессирующие промоторы широкого спектра действия

Можно сказать, что промотор имеет «широкий спектр действия», когда он способствует транскрипции во многих, но необязательно всех, растительных тканях. Например, экспрессирующий промотор широкого спектра действия может способствовать транскрипции функционально связанной последовательности в одном или нескольких побегах, верхушке побега (апексе) и листьях, однако слабо способствует или совсем не способствует транскрипции в таких тканях, как корни или стебли. Другой пример: экспрессирующий промотор широкого спектра действия может способствовать транскрипции функционально связанной последовательности в одном или нескольких стеблях, побегах, верхушке побега (апексе) и листьях, однако может слабо способствовать или совсем не способствовать транскрипции в таких тканях, как репродуктивные ткани цветов и развивающиеся семена. Неограниченные примеры экспрессирующих промоторов широкого спектра действия, которые могут быть включены в конструкции нуклеиновой кислот, приведенных в настоящем документе, включают

промоторы p326, YP0144, YP0190, p13879, YP0050, p32449, 21876, YP0158, YP0214, YP0380, PT0848 и PT0633.

Дополнительные примеры включают промотор вируса мозаики цветной капусты (CaMV) 35S, промотор маннопин-синтазы (MAS), промоторы 1' или 2', полученные из Т-ДНК *Agrobacterium tumefaciens*, промотор вируса мозаики норичника 34S, актиновые промоторы, такие как актиновый промотор риса, и убиквитиновые промоторы, такие как промотор убиквитин-1 кукурузы. В некоторых случаях промотор CaMV 35S исключается из категории экспрессирующих промоторов широкого спектра действия.

ii. Корневые промоторы

Активные промоторы роста корней осуществляют транскрипцию в ткани корня, например, в тканях эндодермы корня, эпидермиса корня или проводящей системы корня. В некоторых вариантах активные корневые промоторы являются преимущественно корневыми промоторами, то есть такими, которые осуществляют транскрипцию только или преимущественно в ткани корня. Преимущественно тканевые промоторы включают промоторы YP0128, YP0275, PT0625, PT0660, PT0683 и PT0758. Другие преимущественно корневые промоторы включают промоторы PT0613, PT0672, PT0688 и PT0837, которые способствуют транскрипции главным образом в ткани корня и в меньшей степени – в семяпочках и/или семенах. Другие примеры преимущественно корневых промоторов включают корнеспецифические субдомены промотора CaMV 35S (Lam *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:7890-7894 (1989)), специфические промоторы корневой клетки, о которых сообщается в работе Conkling *et al.*, *Plant Physiol.*, 93:1203-1211 (1990), и промотор табака RD2.

iii. Промоторы вызревающего эндосперма

В некоторых вариантах могут быть полезны промоторы, которые способствуют транскрипции в вызревающем эндосперме. Транскрипция от промотора вызревающего эндосперма обычно начинается после оплодотворения, происходит главным образом в ткани эндосперма во время развития семени и обычно достигает высшей точки на этапе целлюляризации. Наиболее подходящими являются промоторы, активные, главным образом, в вызревающем эндосперме, хотя иногда могут быть полезны и промоторы, активные в других тканях.

Неограниченные примеры промоторов вызревающего эндосперма, которые могут быть включены в конструкции нуклеиновой кислоты, приведенные в настоящем документе, включают промотор напина, промотор арселина-5, промотор фазеолина (Bustos *et al.*, *Plant Cell*, 1(9):839-853 (1989)), промотор соевых ингибиторов трипсина (Riggs *et al.*, *Plant Cell*, 1(6):609-621 (1989)), промотор АЦП (Baerson *et al.*, *Plant Mol. Biol.*, 22(2):255-267 (1993)), промотор стеароил-ацп десатуразы (Slocombe *et al.*, *Plant Physiol.*, 104(4):167-176 (1994)), промотор сои подгруппы α' β -конглицинина (Chen *et al.*, *Proc. Natl. Acad. Sci. USA*, 83:8560-8564 (1986)), промотор олеозина (Hong *et al.*, *Plant Mol. Biol.*, 34(3):549-555 (1997)) и промоторы зеинов, таких как промотор зеинов 15 kD, промотор зеинов 16 kD, промотор зеинов 19 kD, промотор зеинов 22 kD и промотор зеинов 27 kD. Также пригодны промотор Osgt-1 из гена глютелин-1 риса (Zheng *et al.*, *Mol. Cell Biol.*, 13:5829-5842 (1993)), промотор бета-амилазы и промотор хордеина ячменя. Прочие промоторы вызревающего эндосперма включают промоторы YP0092, PT0676 и PT0708.

iv. Промоторы ткани завязи

Также могут быть полезны промоторы, активные в тканях завязи, таких как стенка семязпочки и мезокарпий, например промотор полигалактуронидазы, промотор TRX банана, промотор актина дыни, YP0396 и PT0623. Примерами промоторов, активных главным образом в семязпочках, являются YP0007, YP0111, YP0092, YP0103, YP0028, YP0121, YP0008, YP0039, YP0115, YP0119, YP0120 и YP0374.

v. Промоторы зародышевого мешка/раннего этапа развития эндосперма

Для достижения экспрессии в зародышевом мешке/на раннем этапе развития эндосперма могут быть использованы регуляторные области, активные в полярных ядрах и/или главной клетке, или в предшественниках полярных ядер, но не в яйцеклетках или предшественниках яйцеклеток. Наиболее подходящими являются промоторы, которые способствуют экспрессии только или преимущественно в полярных ядрах или их предшественниках и/или в главной клетке. Паттерн транскрипции, которая проходит от стадии

полярных ядер до раннего этапа развития эндосперма, может также находиться в промоторах зародышевого мешка/раннего этапа развития эндосперма, хотя обычно транскрипция в значительной степени понижается на поздних этапах развития эндосперма во время и после этапа целлюляризации. С промоторами зародышевого мешка/раннего этапа развития экспрессии в зиготе или в развивающемся эмбрионе эндосперма обычно не происходит.

Подходящими могут быть промоторы, которые получены из следующих генов: *Arabidopsis viviparous-1* (см. GenBank No. U93215); *Arabidopsis atmyc1* (см. Urao, *Plant Mol. Biol.*, 32:571-57 (1996); Conceicao, *Plant*, 5:493-505 (1994)); *Arabidopsis FIE* (GenBank No. AF129516); *Arabidopsis MEA*; *Arabidopsis FIS2* (GenBank No. AF096096) и FIE 1.1 (патент США 6,906,244). Другими подходящими промоторами могут быть те, которые получены из следующих генов: ген кукурузы MAC1 (см. Sheridan, *Genetics*, 142:1009-1020 (1996)); ген кукурузы Cat3 (см. GenBank No. L05934; Abler, *Plant Mol. Biol.*, 22:10131-1038 (1993)). К другим промоторам относятся следующие промоторы *Arabidopsis*: YP0039, YP0101, YP0102, YP0110, YP0117, YP0119, YP0137, DME, YP0285 и YP0212. К другим промоторам, которые могут быть полезны, относятся следующие промоторы риса: p530c10, pOsFIE2-2, pOsMEA, pOsYp102 и pOsYp285.

vi. Промоторы зародышей

Регуляторные области, которые главным образом способствуют транскрипции в зиготных клетках после оплодотворения, могут обеспечивать экспрессию преимущественно на стадии зародыша. Наиболее подходящими являются промоторы, которые главным образом способствуют транскрипции в зародышах на ранних этапах развития до стадии образования сердцевинки, однако также пригодна экспрессия в зародышах на более поздних этапах вызревания. Промоторы преимущественно для стадии зародыша включают промотор липид-переносящих белков (Ltp1) ячменя (*Plant Cell Rep* 20:647-654 (2001)), YP0097, YP0107, YP0088, YP0143, YP0156, PT0650, PT0695, PT0723, PT0838, PT0879 и PT0740.

vii. Промоторы фотосинтетических тканей

Промоторы, активные в фотосинтетических тканях,

осуществляют транскрипцию в зеленых тканях, таких как листья и стебли. Наиболее подходящими являются промоторы, которые способствуют экспрессии только или преимущественно в таких тканях. Примерами таких промоторов являются промоторы рибулозо-1,5-бисфосфаткарбоксилазы (RbcS), такие как промотор RbcS из лиственницы американской (*Larix laricina*), промотор сосны cab6 (Yamamoto et al., *Plant Cell Physiol.*, 35:773-778 (1994)), промотор Cab-1 из пшеницы (Fejes et al., *Plant Mol. Biol.*, 15:921-932 (1990)), промотор CAB-1 из шпината (Lubberstedt et al., *Plant Physiol.*, 104:997-1006 (1994)), промотор cab1R из риса (Luan et al., *Plant Cell*, 4:971-981 (1992)), промотор пируват-ортофосфатдикиназы (PPDK) из кукурузы (Matsuoka et al., *Proc. Natl. Acad. Sci. USA*, 90:9586-9590 (1993)), промотор Lhcb1*2 табака (Cerdan et al., *Plant Mol. Biol.*, 33:245-255 (1997)), промотор *Arabidopsis thaliana* SUC2 симпортера сахарозы-H⁺ (Truernit et al., *Planta*, 196:564-570 (1995)) и промоторы белков тилакоидных мембран из шпината (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Другие промоторы фотосинтетических тканей включают PT0535, PT0668, PT0886, YP0144, YP0380 и PT0585.

viii. Промоторы проводящих тканей

Примерами промоторов, которые имеют высокую или преимущественную активность в сосудистых пучках, являются YP0087, YP0093, YP0108, YP0022 и YP0080. К другим промоторам, которые преимущественно активны в проводящих тканях, относятся промотор GRP 1.8 богатой глицином клеточной стенки белка GRP 1.8 (Keller and Baumgartner, *Plant Cell*, 3(10):1051-1061 (1991)), промотор вируса желтой крапчатости коммелины (CoYMV) (Medberry et al., *Plant Cell*, 4(2):185-192 (1992)) и промотор палочковидного тунгро-вируса риса (RTBV) (Dai et al., *Proc. Natl. Acad. Sci. USA*, 101(2):687-692 (2004)).

ix. Индуцируемые промоторы

Индуцируемые промоторы осуществляют транскрипцию под влиянием внешних стимуляторов, таких как химические вещества, или стимуляторов, связанных с окружающей средой. Например, индуцируемые промоторы могут осуществлять транскрипцию под влиянием гормонов, таких как гиббереллиновая кислота или этилен,

или под влиянием света или засухи. Примерами индуцируемых засухой промоторов являются YP0380, PT0848, YP0381, YP0337, PT0633, YP0374, PT0710, YP0356, YP0385, YP0396, YP0388, YP0384, PT0688, YP0286, YP0377, PD1367 и PD0901. Примерами индуцируемых азотом промоторов являются PT0863, PT0829, PT0665 и PT0886. Примерами индуцируемых тенью промоторов являются PR0924 и PT0678. Примером промотора, индуцируемого солью, служит rd29A (Kasuga *et al.* (1999) *Nature Biotech* 17: 287-291).

х. Базальные промоторы

Базальный промотор является минимальной последовательностью, необходимой для сборки транскрипционного комплекса, требуемого для инициации транскрипции. Базальные промоторы часто включают элемент «ТАТА-бокс», который может быть расположен между примерно 15 и примерно 35 нуклеотидами вышележащей области от сайта инициации транскрипции. Базальные промоторы могут также включать элемент «ССААТ-бокс» (как правило, последовательность ССААТ) и/или последовательность GGGCG, которая может быть расположена между примерно 40 и примерно 200 нуклеотидами, обычно примерно между 60 и примерно 120 нуклеотидами вышележащей области от сайта старта транскрипции.

xi. Промоторы стебля

Промотор стебля может быть специфичным для одной или нескольких тканей стебля или специфичным для стебля и других частей растения. Промоторы стебля могут иметь высокую или преимущественную активность, например, в эпидермисе и кортексе, сосудистом камбии, прокамбии или ксилеме. К примерам промоторов стебля относится YP0018, о котором сообщено в US20060015970, CryIA(b) и CryIA(c) (Braga *et al.* 2003, *Journal of New Seeds* 5:209-221).

xii. Прочие промоторы

Другие классы промоторов включают, в числе прочего, промоторы, преферентные для побегов, каллюса, клеток трихомы, замыкающих клеток, таких как PT0678, клубней, паренхимных клеток и для старения. Также могут быть полезны промоторы, обозначенные YP0086, YP0188, YP0263, PT0758, PT0743, PT0829, YP0119 и YP0096,

как описано в вышеуказанных заявках на патент.

xiii. Прочие регуляторные области

В конструкции нуклеиновой кислоты, описанные в настоящем документе, может быть включена нетранслируемая область 5' (UTR). Нетранслируемая область 5' транскрибируется, но не транслируется лежит между сайтом старта транскрипции и кодоном инициации трансляции и может включать +1 нуклеотид. Нетранслируемая область 3' может находиться между терминирующим трансляцию кодоном и концом транскрипции. Нетранслируемые области могут выполнять специфические функции, такие как увеличение стабильности мРНК или ослабление трансляции. К примерам нетранслируемых областей 3' относятся в числе прочего сигналы полиаденилирования и терминирующие транскрипцию последовательности, *например* терминирующую нопалинсентазу последовательность.

Следует понимать, что в рекомбинантном полинуклеотиде может присутствовать более одной регуляторной области, *например* интроны, энхансеры, вышележащие области активации, терминаторы транскрипции и индуцируемые элементы. Таким образом, к примеру, более чем одна регуляторная область может быть функционально связана с последовательностью полинуклеотида, который кодирует моделирующий биомассу полипептид.

Регуляторные области, такие как промоторы эндогенных генов, могут быть получены путем химического синтеза или путем субклонирования из геномной ДНК, которая включает такую регуляторную область. Нуклеиновая кислота, содержащая такую регуляторную область, может также включать примыкающие последовательности, которые содержат сайты рестрикционных ферментов, облегчающих последующую манипуляцию.

IV. Трансгенные растения и растительные клетки

A. Трансформация

В изобретении также представлены трансгенные растительные клетки и растения, содержащие как минимум одну рекомбинантную конструкцию нуклеиновой кислоты, описанную в настоящем документе. Растение или растительная клетка могут быть трансформированы путем интеграции конструкции в их геном, *то*

есть они могут быть устойчиво трансформированными. Устойчиво трансформированные клетки обычно удерживают интродуцированную нуклеиновую кислоту при каждом делении клетки. Растение или растительная клетка могут быть также трансформированы временно таким образом, чтобы конструкция не интегрировалась в их геном. Временно трансформированные клетки обычно полностью или частично теряют интродуцированную конструкцию нуклеиновой кислоты при каждом делении клетки, так что после значительного количества делений клеток интродуцированную нуклеиновую кислоту обнаружить в дочерних клетках невозможно. И временно трансформированные, и устойчиво трансформированные трансгенные растения и растительные клетки могут быть полезны при использовании методов, описанных в настоящем документе.

Трансгенные растительные клетки, которые используются в методах, описанных в настоящем документе, могут составлять часть растения или растение целиком. Такие растения могут выращиваться пригодным для рассматриваемых видов способом – в ростовой камере, теплице или в поле. Трансгенные растения при желании можно разводить для определенных целей, например для интродуцирования рекомбинантных нуклеиновых кислот в другие линии, для переноса рекомбинантной нуклеиновой кислоты в другие виды или для дальнейшей селекции желаемых признаков. Или же трансгенные растения можно размножать вегетативно, применительно к видам, восприимчивым к таким технологиям. Для целей настоящего документа трансгенное растение также относится к потомству исходного трансгенного растения при условии, что потомство наследует трансген. Семена, полученные от трансгенного растения, могут выращиваться и затем самовоспроизводиться (или скрещиваться с сортами других линий и самовоспроизводиться) для получения семян, гомозиготных по конструкции нуклеиновой кислоты.

Трансгенные растения можно выращивать в суспензионной культуре, или ткани, или культуре органов. Для целей настоящего изобретения могут быть использованы методы выращивания тканевой культуры на жидкой и/или твердой питательной основе. При использовании твердой среды трансгенные растительные клетки

можно помещать непосредственно в среду или можно разместить на фильтре, который затем помещают в среду. При использовании жидкой среды трансгенные растительные клетки можно помещать на флотационное устройство, например на пористую мембрану, которая находится в контакте с жидкой средой. Твердой средой может быть, например, агар с питательной средой Мурасиге-Скуга (MS) с подходящей концентрацией ауксина, например 2, 4-дихлорфеноксиуксусной кислоты (2,4-D), и подходящей концентрацией цитокинина, например кинетина.

При использовании временно трансформированных растительных клеток последовательность репортера, которая кодирует полипептид репортера, имеющий активность репортера, может быть включена в процедуру трансформации, и испытание активности или экспрессии репортера можно выполнить в подходящее время после трансформации. Подходящее время для проведения испытания может обычно составлять примерно 1-21 день после трансформации, например примерно 1-14 дней, примерно 1-7 дней или примерно 1-3 дня. Промежуточные испытания особенно удобны для проведения экспресс-анализов у различных видов или для подтверждения экспрессии гетерологичного модулирующего биомассу полипептида, экспрессия которого в определенных клетках-реципиентах не была подтверждена ранее.

Способы интродукции нуклеиновых кислот в однодольные и двудольные растения в данной отрасли знаний известны и включают, помимо прочего, агробактериальную трансформацию, трансформацию с помощью вирусных векторов, трансформацию с помощью электропорации и генных пушек, например, патенты США 5,538,880; 5,204,253; 6,329,571 и 6,013,863. Если для трансформации в качестве ткани-реципиента используется клетка или культура ткани, выращенная в питательной среде, при желании растения могут быть регенерированы из трансформированных культур способами, которые известны специалистам в данной области.

Б. Скрининг/отбор

Можно провести скрининг популяции трансгенных растений и/или отбор в ней тех представителей популяции, у которых признаки или фенотип передаются путем экспрессии трансгена.

Например, можно провести скрининг популяции потомков отдельного объекта трансформации по тем растениям, которые показывают необходимый уровень экспрессии модулирующего биомассу полипептида или нуклеиновой кислоты. Для определения уровней экспрессии можно использовать физические и биохимические методы. Они включают саузерн-блоттинг или ПЦР-амплификацию для обнаружения полинуклеотида; нозерн-блоттинг, защиту от S1 РНКазы, метод удлинения затравки или амплификацию ОТ-ПЦР для определения транскриптов РНК; ферментные анализы для определения активности ферментов или рибозимов полипептидов и полинуклеотидов; электрофорез белкового геля, вестерн-блоттинг, иммунопреципитацию и иммуноферментный анализ для определения полипептидов. Другие способы, такие как гибридизация *in situ*, окрашивание ферментов и иммуноокрашивание могут быть также использованы для определения наличия или экспрессии полипептидов и/или полинуклеотидов. Способы выполнения всех упомянутых приемов известны. В качестве варианта можно провести скрининг популяции растений с независимыми событиями трансформации по тем растениям, которые показывают необходимый признак, такой как модулированный уровень биомассы. Отбор и/или скрининг можно проводить по одному и/или нескольким поколениям и/или по нескольким географическим местам. В некоторых случаях можно выращивать трансгенные растения и производить их отбор в условиях, которые индуцируют желаемый фенотип или же необходимы для получения желаемого фенотипа у трансгенного растения. Кроме того, отбор и/или скрининг можно применять во время определенной стадии развития, на которой ожидается, что растение проявит фенотип. Отбор и/или скрининг можно проводить для отбора трансгенных растений, имеющих статистически достоверную разницу в уровне биомассы по сравнению с контрольным растением, у которого трансген отсутствует. Отобранные или прошедшие скрининг трансгенные растения имеют измененный фенотип по сравнению с соответствующим контрольным растением, как описано в разделе «Фенотипы трансгенных растений» настоящего документа.

В. Виды растений

Полинуклеотиды и векторы, описанные в настоящем документе,

можно использовать для трансформации ряда однодольных и двудольных растений и систем растительных клеток, в том числе видов одного из следующих семейств: *Acanthaceae*, *Alliaceae*, *Alstroemeriaceae*, *Amaryllidaceae*, *Apocynaceae*, *Arecaceae*, *Asteraceae*, *Berberidaceae*, *Bixaceae*, *Brassicaceae*, *Bromeliaceae*, *Cannabaceae*, *Caryophyllaceae*, *Cephalotaxaceae*, *Chenopodiaceae*, *Colchicaceae*, *Cucurbitaceae*, *Dioscoreaceae*, *Ephedraceae*, *Erythroxylaceae*, *Euphorbiaceae*, *Fabaceae*, *Lamiaceae*, *Linaceae*, *Lycopodiaceae*, *Malvaceae*, *Melanthiaceae*, *Musaceae*, *Myrtaceae*, *Nyssaceae*, *Papaveraceae*, *Pinaceae*, *Plantaginaceae*, *Poaceae*, *Rosaceae*, *Rubiaceae*, *Salicaceae*, *Sapindaceae*, *Solanaceae*, *Taxaceae*, *Theaceae* или *Vitaceae*.

Подходящие виды могут включать представителей родов *Abelmoschus*, *Abies*, *Acer*, *Agrostis*, *Allium*, *Alstroemeria*, *Ananas*, *Andrographis*, *Andropogon*, *Artemisia*, *Arundo*, *Atropa*, *Berberis*, *Beta*, *Bixa*, *Brassica*, *Calendula*, *Camellia*, *Camptotheca*, *Cannabis*, *Capsicum*, *Carthamus*, *Catharanthus*, *Cephalotaxus*, *Chrysanthemum*, *Cinchona*, *Citrullus*, *Coffea*, *Colchicum*, *Coleus*, *Cucumis*, *Cucurbita*, *Cynodon*, *Datura*, *Dianthus*, *Digitalis*, *Dioscorea*, *Elaeis*, *Ephedra*, *Erianthus*, *Erythroxylum*, *Eucalyptus*, *Festuca*, *Fragaria*, *Galanthus*, *Glycine*, *Gossypium*, *Helianthus*, *Hevea*, *Hordeum*, *Hyoscyamus*, *Jatropha*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Lycopodium*, *Manihot*, *Medicago*, *Mentha*, *Miscanthus*, *Musa*, *Nicotiana*, *Oryza*, *Panicum*, *Papaver*, *Parthenium*, *Pennisetum*, *Petunia*, *Phalaris*, *Phleum*, *Pinus*, *Poa*, *Poinsettia*, *Populus*, *Rauwolfia*, *Ricinus*, *Rosa*, *Saccharum*, *Salix*, *Sanguinaria*, *Scopolia*, *Secale*, *Solanum*, *Sorghum*, *Spartina*, *Spinacea*, *Tanacetum*, *Taxus*, *Theobroma*, *Triticosecale*, *Triticum*, *Uniola*, *Veratrum*, *Vinca*, *Vitis* и *Zea*.

Подходящие виды включают *Panicum* spp., *Sorghum* spp., *Miscanthus* spp., *Saccharum* spp., *Erianthus* spp., *Populus* spp., *Andropogon gerardii* (бородач), *Pennisetum purpureum* (слоновая трава), *Phalaris arundinacea* (канареечник трубковидный), *Cynodon dactylon* (бермудская трава), *Festuca arundinacea* (овсяница высокая), *Spartina pectinata* (спартина гребенчатая), *Medicago sativa* (люцерна), *Arundo donax* (арундо тростниковый), *Secale*

cereale (рожь), *Salix* spp. (ива), *Eucalyptus* spp. (эвкалипт), *Triticosecale* (тритикум - гибрид пшеницы и ржи) и бамбук.

Подходящие виды также включают *Helianthus annuus* (подсолнечник), *Carthamus tinctorius* (сафлор), *Jatropha curcas* (ятрофа), *Ricinus communis* (клещевина обыкновенная), *Elaeis guineensis* (гвинейская масличная пальма), *Linum usitatissimum* (лен посевной) и *Brassica juncea*.

Подходящие виды также включают *Beta vulgaris* (сахарная свекла) и *Manihot esculenta* (маниок съедобный).

Подходящие виды также включают *Lycopersicon esculentum* (помидор), *Lactuca sativa* (салат), *Musa paradisiaca* (банан райский), *Solanum tuberosum* (картофель), *Brassica oleracea* (капуста спаржевая, цветная капуста, капуста брюссельская), *Camellia sinensis* (чай китайский), *Fragaria ananassa* (земляника садовая), *Theobroma cacao* (дерево какао), *Coffea arabica* (кофейное дерево аравийское), *Vitis vinifera* (виноград винный), *Ananas comosus* (ананас), *Capsicum annum* (перец горький и сладкий), *Allium cepa* (лук репчатый), *Cucumis melo* (дыня), *Cucumis sativus* (огурец), *Cucurbita maxima* (тыква крупноплодная), *Cucurbita moschata* (тыква мускатная), *Spinacea oleracea* (шпинат огородный), *Citrullus lanatus* (арбуз обыкновенный), *Abelmoschus esculentus* (окра) и *Solanum melongena* (баклажан).

Подходящие виды также включают *Papaver somniferum* (мак снотворный), *Papaver orientale*, *Taxus baccata*, *Taxus brevifolia*, *Artemisia annua*, *Cannabis sativa*, *Camptotheca acuminata*, *Catharanthus roseus*, *Vinca rosea*, *Cinchona officinalis*, *Colchicum autumnale*, *Veratrum californica*, *Digitalis lanata*, *Digitalis purpurea*, *Dioscorea* spp., *Andrographis paniculata*, *Atropa belladonna*, *Datura stomonium*, *Berberis* spp., *Cephalotaxus* spp., *Ephedra sinica*, *Ephedra* spp., *Erythroxylum coca*, *Galanthus wornorii*, *Scopolia* spp., *Lycopodium serratum* (*Huperzia serrata*), *Lycopodium* spp., *Rauwolfia serpentina*, *Rauwolfia* spp., *Sanguinaria canadensis*, *Hyoscyamus* spp., *Calendula officinalis*, *Chrysanthemum parthenium*, *Coleus forskohlii* и *Tanacetum parthenium*.

Подходящие виды также включают *Parthenium argentatum* (гваюла серебристая), *Nevea* spp. (гевея бразильская), *Mentha spicata* (мята колосовая), *Mentha piperita* (мята перечная), *Bixa orellana* и *Alstroemeria* spp.

Подходящие виды также включают *Rosa* spp. (роза), *Dianthus caryophyllus* (гвоздика садовая), *Petunia* spp. (петуния) и *Poinsettia pulcherrima* (пуансеттия).

Подходящие виды также включают *Nicotiana tabacum* (табак), *Lupinus albus* (люпин белый), *Uniola paniculata* (униола метельчатая), полевица (*Agrostis* spp.), *Populus tremuloides* (тополь осинообразный), *Pinus* spp. (сосна), *Abies* spp. (пихта), *Acer* spp. (клен), *Hordeum vulgare* (ячмень посевной), *Poa pratensis* (мятлик луговой), *Lolium* spp. (плевел) и *Phleum pratense* (тимофеевка луговая).

В некоторых вариантах подходящими видами могут быть дикие, сорные или культивированные виды *Pennisetum*, такие как *Pennisetum alopecuroides*, *Pennisetum arnhemicum*, *Pennisetum caffrum*, *Pennisetum clandestinum*, *Pennisetum divisum*, *Pennisetum glaucum*, *Pennisetum latifolium*, *Pennisetum macrostachyum*, *Pennisetum macrourum*, *Pennisetum orientale*, *Pennisetum pedicellatum*, *Pennisetum polystachion*, *Pennisetum polystachion* ssp. *Setosum*, *Pennisetum purpureum*, *Pennisetum setaceum*, *Pennisetum subangustum*, *Pennisetum typhoides*, *Pennisetum villosum* или их гибриды (например, *Pennisetum purpureum* x *Pennisetum typhoidum*).

В некоторых вариантах подходящими видами могут быть дикие, сорные или культивированные виды и/или разновидности *Miscanthus*, такие как *Miscanthus* x *giganteus*, *Miscanthus sinensis*, *Miscanthus* x *ogiformis*, *Miscanthus floridulus*, *Miscanthus transmorrisonensis*, *Miscanthus oligostachyus*, *Miscanthus nepalensis*, *Miscanthus sacchariflorus*, *Miscanthus* x *giganteus* "Amuri", *Miscanthus* x *giganteus* "Nagara", *Miscanthus* x *giganteus* "Illinois", *Miscanthus sinensis* var. "Goliath", *Miscanthus sinensis* var. "Roland", *Miscanthus sinensis* var. "Africa", *Miscanthus sinensis* var. "Fern Osten", *Miscanthus sinensis* var. *gracillimus*, *Miscanthus sinensis* var. *variegates*, *Miscanthus*

sinensis var. *purpurascens*, *Miscanthus sinensis* var. "Malepartus", *Miscanthus sacchariflorus* var. "Robusta", *Miscanthus sinensis* var. "Silberfedher" (он же Silver Feather), *Miscanthus transmorrissonensis*, *Miscanthus condensatus*, *Miscanthus yakushimanum*, *Miscanthus* var. "Alexander", *Miscanthus* var. "Adagio", *Miscanthus* var. "Autumn Light", *Miscanthus* var. "Cabaret", *Miscanthus* var. "Condensatus", *Miscanthus* var. "Cosmopolitan", *Miscanthus* var. "Dixieland", *Miscanthus* var. "Gilded Tower" (патент США № PP14,743), *Miscanthus* var. "Gold Bar" (патент США № PP15,193), *Miscanthus* var. "Gracillimus", *Miscanthus* var. "Graziella", *Miscanthus* var. "Grosse Fontaine", *Miscanthus* var. "Hinjo aka Little Nicky" ™, *Miscanthus* var. "Juli", *Miscanthus* var. "Kaskade", *Miscanthus* var. "Kirk Alexander", *Miscanthus* var. "Kleine Fontaine", *Miscanthus* var. "Kleine Silberspinne" (он же "Little Silver Spider"), *Miscanthus* var. "Little Kitten", *Miscanthus* var. "Little Zebra" (патент США № PP13,008), *Miscanthus* var. "Lottum", *Miscanthus* var. "Malepartus", *Miscanthus* var. "Morning Light", *Miscanthus* var. "Mysterious Maiden" (патент США № PP16,176), *Miscanthus* var. "Nippon", *Miscanthus* var. "November Sunset", *Miscanthus* var. "Parachute", *Miscanthus* var. "Positano", *Miscanthus* var. "Puenktchen" (он же "Little Dot"), *Miscanthus* var. "Rigoletto", *Miscanthus* var. "Sarabande", *Miscanthus* var. "Silberpfeil" (он же Silver Arrow), *Miscanthus* var. "Silverstripe", *Miscanthus* var. "Super Stripe" (патент США № PP18,161), *Miscanthus* var. "Strictus" или *Miscanthus* var. "Zebrinus".

В некоторых вариантах подходящими видами могут быть дикие, сорные или культивированные виды и/или разновидности сорго, такие как *Sorghum alnum*, *Sorghum amplum*, *Sorghum angustum*, *Sorghum arundinaceum*, *Sorghum bicolor* (такие как двуцветное, гвинейское, хвостатое, кафрское и зерновое), *Sorghum brachypodium*, *Sorghum bulbosum*, *Sorghum burmahicum*, *Sorghum controversum*, *Sorghum drummondii*, *Sorghum ecarinatum*, *Sorghum exstans*, *Sorghum grande*, *Sorghum halepense*, *Sorghum interjectum*, *Sorghum intrans*, *Sorghum laxiflorum*, *Sorghum leiocladum*, *Sorghum macrospermum*, *Sorghum matarankense*, *Sorghum miliaceum*, *Sorghum*

nigrum, *Sorghum nitidum*, *Sorghum plumosum*, *Sorghum propinquum*, *Sorghum purpureosericeum*, *Sorghum stipoides*, *Sorghum sudanense*, *Sorghum timorense*, *Sorghum trichocladum*, *Sorghum versicolor*, *Sorghum virgatum*, *Sorghum vulgare* или гибриды, такие как *Sorghum* × *almum*, *Sorghum* × суданская трава или *Sorghum* × *drummondii*.

Таким образом, методы и составы можно использовать в широком диапазоне видов растений, в том числе видов из родов двудольных растений *Brassica*, *Carthamus*, *Glycine*, *Gossypium*, *Helianthus*, *Jatropha*, *Parthenium*, *Populus* и *Ricinus*; родов однодольных растений *Elaeis*, *Festuca*, *Hordeum*, *Lolium*, *Oryza*, *Panicum*, *Pennisetum*, *Phleum*, *Poa*, *Saccharum*, *Secale*, *Sorghum*, *Triticosecale*, *Triticum* и *Zea*. В некоторых вариантах растение является представителем видов *Panicum virgatum* (просо), *Sorghum bicolor* (сорго, суданская трава), *Miscanthus giganteus* (мискантус), *Saccharum* sp. (сахарный тростник), *Populus balsamifera* (тополь бальзамический), *Zea mays* (маис), *Glycine max* (соя), *Brassica napus* (канола), *Triticum aestivum* (пшеница), *Gossypium hirsutum* (хлопчатник обыкновенный), *Oryza sativa* (рис), *Helianthus annuus* (подсолнечник однолетний), *Medicago sativa* (люцерна), *Beta vulgaris* (сахарная свекла) или *Pennisetum glaucum* (просо африканское).

В некоторых вариантах полинуклеотиды и векторы, описанные в настоящем документе, можно использовать для трансформации ряда однодольных и двудольных растений и систем растительных клеток в тех случаях, когда такие растения являются гибридами различных видов или разновидностей конкретного вида (например, *Saccharum* sp. × *Miscanthus* sp., *Sorghum* sp. × *Miscanthus* sp., например: *Panicum virgatum* × *Panicum amarum*, *Panicum virgatum* × *Panicum amarulum* и *Pennisetum purpureum* × *Pennisetum typhoidum*).

Г. Фенотипы трансгенных растений

В некоторых вариантах растение, у которого модулируется экспрессия модулирующего биомассу полипептида, может иметь повышенный уровень биомассы. Например, модулирующий биомассу полипептид, описанный в настоящем документе, может быть экспрессирован в трансгенном растении, что приводит к увеличению

уровня вегетативной ткани. Уровень биомассы может быть увеличен минимум на 2 процента, *например* на 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60 или более 60 процентов по сравнению с соответствующим уровнем биомассы контрольного растения, которое не экспрессирует трансген. В некоторых вариантах растение, у которого модулируется экспрессия модулирующего биомассу полипептида, может иметь пониженный уровень семенной продуктивности. Уровень может быть уменьшен минимум на 2 процента, *например* на 2, 3, 4, 5, 10, 15, 20, 25, 30, 35 или более 35 процентов по сравнению с соответствующим уровнем семенной продуктивности контрольного растения, которое не экспрессирует трансген.

Увеличение семенной продуктивности у таких растений может обеспечить повышение доступности питания в географических регионах, где потребление пищевых продуктов часто недостаточно, или обеспечить производство биотоплива. В некоторых вариантах уменьшение биомассы у таких растений может быть полезным, если вегетативные ткани не являются основной частью растения, которую получают для потребления человеком или животными (*например*, сбор семян).

В некоторых вариантах растение, у которого модулируется экспрессия модулирующего биомассу полипептида, может иметь повышенный или пониженный уровень биомассы в одной или нескольких тканях, *например* в вегетативных тканях, репродуктивных тканях или тканях корня. *Например*, уровень биомассы может быть увеличен минимум на 2 процента, *например* на 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60 или более 60 процентов по сравнению с соответствующим уровнем биомассы контрольного растения, которое не экспрессирует трансген. В некоторых вариантах растение, у которого модулируется экспрессия модулирующего биомассу полипептида, может иметь пониженный уровень биомассы в одной или нескольких растительных тканей. Уровень биомассы может быть уменьшен минимум на 2 процента, *например* на 2, 3, 4, 5, 10, 15, 20, 25, 30, 35 или более 35 процентов по сравнению с соответствующим уровнем биомассы

контрольного растения, которое не экспрессирует трансген.

Увеличение биомассы у таких растений может обеспечить улучшение качества пищи или повышенную энергетическую продуктивность. Уменьшение биомассы может обеспечить более эффективное распределение питательных веществ по частям растений, которые выращивают для потребления человеком или животными.

Как правило, разница в количестве биомассы у трансгенного растения или трансгенной клетки по сравнению с контрольным растением или клеткой считается статистически достоверной при значении $p \leq 0,05$ при применении параметрических или непараметрических методов статистической оценки, таких как проверка на соответствие по критерию хи-квадрат, t-критерий Стьюдента, критерий Манна-Уитни или F-критерий Фишера. В некоторых вариантах разница в количестве биомассы является статистически достоверной при $p < 0,01$, $p < 0,005$ или $p < 0,001$. Статистически достоверная разница, например, в количестве биомассы у трансгенного растения по сравнению с ее количеством у контрольного растения показывает, что присутствие рекомбинантной нуклеиновой кислоты в трансгенном растении приводит к изменению уровня биомассы.

Фенотип трансгенного растения оценивается в сравнении с контрольным растением. Считается, что растение «не экспрессирует» полипептид, когда в растении обнаруживается менее 10%, например менее 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0,5%, 0,1%, 0,01% или 0,001% количества полипептида или мРНК, кодирующей полипептид, обнаруживаемый в рассматриваемом растении. Экспрессию можно оценить такими методами, как, например, ОТ-ПЦР, нозерн-блоттинг, защита от S1 РНКазы, удлинение затравки, вестерн-блоттинг, электрофорез белкового геля, иммунопреципитация, иммуноферментный анализ, иммунопреципитация хроматина и масс-спектрометрия. Следует отметить, что, если полипептид экспрессируется под контролем промотора, преимущественно активного в определенных тканях, или промотора широкого спектра действия, экспрессию можно оценить во всем растении или в выбранной ткани. Аналогичным образом, если

полипептид экспрессируется в определенное время, например на определенном этапе развития или во время индукции, экспрессию можно оценить выборочно в требуемый период времени.

Биомасса может включать собираемые ткани растений, такие как листья, стебли и репродуктивные структуры, или все растительные ткани, такие как листья, стебли, корни и репродуктивные структуры. В некоторых вариантах биомасса включает только надземные части растений. В некоторых вариантах биомасса включает только стеблевые части растений. В некоторых вариантах биомасса включает только надземные части растений, за исключением соцветий и семенных частей растений. Биомассу можно измерять способами, описанными в разделе «Примеры». Биомассу количественно можно выразить в виде выхода сухого вещества, что является массой полученной биомассы (обычно выраженной в т/акр), если из массы свежей продукции вычесть воду. Выход сухого вещества (DMY) вычисляется с использованием массы свежей продукции (FMW) и измерения весового процента влаги (M) по следующему уравнению: $DMY = ((100 - M) / 100) * FMW$. Биомассу количественно можно выразить в виде массы свежей продукции, то есть массы биомассы (обычно выраженной в т/акр), полученной непосредственно после выращивания, которая включает массу влаги.

V. Модификация эндогенных нуклеиновых кислот, которые кодируют модулирующие биомассу полипептиды

В настоящем документе также представлены растительные клетки и растения, в которых эндогенная модулирующая биомассу нуклеиновая кислота, описанная в настоящем документе, была модифицирована (например, были модифицированы регуляторная область, интрон или кодирующая область модулирующей биомассу нуклеиновой кислоты). Биомасса таких растений изменена по сравнению с соответствующим уровнем биомассы контрольного растения, в котором эндогенная нуклеиновая кислота не модифицирована. Такие растения рассматриваются в настоящем документе как модифицированные растения, и они могут быть использованы, например, для получения увеличенного количества биомассы.

Эндогенная нуклеиновая кислота может быть модифицирована

методом гомологической рекомбинации. Например, специфические для последовательности эндонуклеазы (например, нуклеазы белкового домена «цинковые пальцы» (ZFN)) и мегануклеазы можно использовать для стимуляции гомологических рекомбинаций в эндогенных генах растений. См., например, Townsend *et al.*, *Nature* 459:442-445 (2009); Tovkach *et al.*, *Plant J.*, 57:747-757 (2009); и Lloyd *et al.*, *Proc. Natl. Acad. Sci. USA*, 102:2232-2237 (2005). В частности, нуклеазы белкового домена «цинковые пальцы», сконструированные для создания парных разрывов двухцепочечной ДНК в конкретных локусах, можно использовать для направленных изменений последовательностей в эндогенных генах растений. Например, эндогенный ген растения может быть заменен вариантом с одной или несколькими мутациями (например, полученными с использованием сайт-направленного мутагенеза или направленной эволюции).

В некоторых вариантах эндогенные нуклеиновые кислоты можно модифицировать путем метилирования или деметилирования, при которых изменяется экспрессия модифицированной эндогенной нуклеиновой кислоты. Например, можно использовать двухцепочечную РНК для активации экспрессии гена путем направленного воздействия на некодирующие регуляторные области в промоторах генов. См. Shibuya *et al.*, *Proc Natl Acad Sci USA*, 106(5): 1660-1665 (2009); и Li *et al.*, *Proc Natl Acad Sci USA*, 103(46):17337-42 (2006).

В некоторых вариантах эндогенные нуклеиновые кислоты можно модифицировать путем введения меток активации. Например, для активации эндогенного гена можно использовать вектор, содержащий несколько копий элемента энхансера от конститутивно активного промотора гена вируса мозаики цветной капусты (CaMV) 35S. См. Weigel *et al.*, *Plant Physiology*, 122:1003-1013 (2000).

В некоторых вариантах эндогенные нуклеиновые кислоты можно модифицировать путем интродукции сконструированного фактора транскрипции, участвующего в активации/репрессии (например, фактора транскрипции белка «цинковые пальцы» или ZFP TF. См., например, [sangamo.com/tech/tech_plat_over.html#whatarezfp](http://www.sangamo.com/tech/tech_plat_over.html#whatarezfp)).

в
Интернете

Сконструированный фактор транскрипции, участвующий в активации/репрессии (например, ZFP TF), может активировать, репрессировать или запускать экспрессию целевого эндогенного гена, связанного с ростом биомассы, путем связывания особым образом эндогенного гена с областью промотора или кодирующей областью.

В некоторых вариантах эндогенные нуклеиновые кислоты можно модифицировать путем мутагенеза. Интродукцию генетических мутаций в регенерируемых растительных тканях можно произвести с помощью одного или более мутагенных агентов. Подходящими мутагенными агентами являются, например, этилметансульфонат (EMS), N-нитрозо-N-этилмочевина (ENU), метил-нитрозогуанидин (MNNG), бромид этидия, диэпоксидбутан, ионизирующее излучение, рентгеновское излучение, ультрафиолетовое излучение и другие мутагены, известные в данной отрасли знаний. Подходящие типы мутаций включают, например, вставки или делеции нуклеотидов и транзиции или трансверсии в последовательности эндогенных нуклеиновых кислот. В одном варианте для получения растений, имеющих модифицированную эндогенную нуклеиновую кислоту, можно использовать TILLING-метод (метод индуцированного мутагенеза). TILLING-метод объединяет мутагенез высокой интенсивности с методами высокопроизводительного скрининга. См. например, McCallum *et al.*, *Nat Biotechnol* 18: 455-457 (2000); reviewed by Stemple, *Nat Rev Genet* 5(2):145-50 (2004).

В некоторых вариантах эндогенную нуклеиновую кислоту можно модифицировать путем подавления активности гена. См., например, раздел настоящего документа «Ингибирование экспрессии модулирующего биомассу полипептида».

Можно провести скрининг популяции растений и/или отбор в ней тех представителей популяции, которые имеют модифицированную нуклеиновую кислоту. Можно провести скрининг популяции растений и/или отбор в ней тех представителей популяции, у которых признак или фенотип передаются путем экспрессии модифицированной нуклеиновой кислоты. В качестве варианта можно провести скрининг популяции тех растений, которые проявляют желаемый признак, такой как модулированный уровень биомассы. Например, можно

провести скрининг популяции потомков тех растений, которые имеют желаемый уровень экспрессии модулирующего биомассу полипептида или нуклеиновой кислоты. Для определения модифицированных нуклеиновых кислот и/или уровней экспрессии можно использовать физические и биохимические методы, как указывалось при описании трансгенных растений. Отбор и/или скрининг можно проводить по одному и/или нескольким поколениям, а также по одному и/или нескольким географическим местам. В некоторых случаях можно выращивать растения и производить их отбор в условиях, индуцирующих желаемый фенотип или необходимых для получения желаемого фенотипа у модифицированного растения. Кроме того, отбор и/или скрининг можно применять во время определенной стадии развития, на которой у растения ожидается проявление фенотипа. Отбор и/или скрининг можно проводить для выбора модифицированных растений, имеющих статистически достоверную разницу в уровне биомассы по сравнению с контрольным растением, у которого нуклеиновая кислота не была модифицирована. Отобранные или прошедшие скрининг модифицированные растения имеют измененный фенотип по сравнению с соответствующим контрольным растением, как описано в разделе «Фенотипы трансгенных растений» настоящего документа.

Несмотря на то, что растение или растительная клетка, у которых эндогенная модулирующая биомассу нуклеиновая кислота была модифицирована, не являются трансгенными для данной конкретной нуклеиновой кислоты, следует понимать, что такое растение или такая растительная клетка может содержать трансгены. Например, модифицированное растение может содержать трансген для других признаков, таких как устойчивость к гербицидам или насекомым. Другой пример: модифицированное растение может содержать один или несколько трансгенов, что в сочетании с модификациями одной или нескольких эндогенных нуклеиновых кислот проявляется в увеличении биомассы.

Как и в случае с трансгенными растительными клетками, модифицированные растительные клетки могут составлять часть растения или растение целиком. Такие растения можно выращивать таким же способом, который описан для трансгенных растений, и

можно разводить и размножать таким же способом, который описан для трансгенных растений.

VI. Селекция растений

Генетический полиморфизм, полезный при использовании таких методов, включает простые повторения последовательностей (SSR, или микросателлиты), быструю амплификацию полиморфной ДНК (RAPD), однонуклеотидные полиморфизмы (SNP), полиморфизмы длины амплификационных фрагментов (ПДАФ) и полиморфизмы длины рестрикционных фрагментов (ПДРФ). Полиформизмы с простыми повторениями последовательностей (SSR) можно идентифицировать, например, с помощью сайт-специфических зондов и амплификации матричной ДНК от отдельных представителей популяции, представляющей интерес, путем ПЦР. Например, метод ПЦР можно использовать для ферментативной амплификации генетического маркера, связанного с нуклеотидной последовательностью, которая передает конкретный признак (например, нуклеотидные последовательности, описанные в настоящем документе). ПЦР можно использовать для амплификации конкретных последовательностей ДНК и РНК, в том числе последовательностей всей геномной ДНК или всей клеточной РНК. При использовании РНК в качестве матрицы для синтеза нитей комплементарной ДНК (кДНК) можно использовать обратную транскриптазу. Различные методы ПЦР описаны, например, в праймере для ПЦР: A Laboratory Manual, Dieffenbach and Dveksler, eds., Cold Spring Harbor Laboratory Press, 1995.

Как правило, для конструирования олигонуклеотидных праймеров, которые в последовательности идентичны или подобны противоположащим цепям амплифицируемой матрицы, используется информация от полинуклеотидов, примыкающих к представляющей интерес области, или из участков, находящихся за ее пределами. Праймеры обычно имеют длину 14-40 нуклеотидов, однако могут быть длиной от 10 нуклеотидов до нескольких сотен нуклеотидов. Для отделения двойной цепи матрица или амплифицированная ДНК многократно денатурируется при высокой температуре, затем охлаждается, чтобы можно было произвести отжиг праймеров и протянуть нуклеотидные последовательности через микросателлиты,

и в результате нарабатывается соответствующая ДНК, достаточная для определения продуктов ПЦР. Если зонды в популяции примыкают к маркеру SSR, вырабатываются продукты ПЦР различных размеров. См., например, патент США 5,766,847.

Произвести качественный и количественный анализ продуктов ПЦР можно несколькими способами. Например, продукты ПЦР можно пометить флуоресцирующей молекулой (например, PicoGreen® или OliGreen®) или определить в растворе с помощью спектрофотометрии или капиллярного электрофореза. В некоторых случаях продукты ПЦР можно отделить в гелевой матрице (например, агарозе или полиакриламиде) путем электрофореза, а фракционированные по размеру полосы белков, содержащие продукты ПЦР, можно сделать видимыми путем окрашивания нуклеиновых кислот. Подходящие красители могут флуоресцировать в ультрафиолетовом свете (например, бромид этидия, GR Safe, SYBR® Green или SYBR® Gold). Результаты можно сделать видимыми с помощью просвечивания или эпи-иллюминации, а изображение флуоресцирующего паттерна можно получить, например, с помощью фотоаппарата или сканера. Изображение можно обработать и проанализировать с помощью специального программного обеспечения (например, программы ImageJ), чтобы произвести измерение и сравнение интенсивности представляющей интерес полосы с калибровочной пробой, загруженной в тот же гель.

Или же полиморфизмы с простыми повторениями последовательностей (SSR) можно идентифицировать, используя продукт (ы) ПЦР в качестве зонда в сопоставлении с саузерн-блотами от разных представителей в популяции. См. U.N. Refseth et al., (1997) *Electrophoresis* 18: 1519. Если говорить коротко, продукты ПЦР разделяются по длине путем гель-электрофореза и переносятся на мембрану. Микросателлитные ДНК-зонды, такие как олигонуклеотид, помеченные радиоактивными, флуоресцирующими или хромогенными молекулами, подаются на мембрану и гибридизируются, чтобы связать продукты ПЦР с комплементарной нуклеотидной последовательностью. Паттерн гибридизации можно сделать видимым, например, с помощью авторадиографии или проявления цвета на мембране.

В некоторых случаях можно определить количество продуктов ПЦР, используя термоциклер с системой детекции в режиме реального времени.

Например, при количественном ПЦР-анализе в режиме реального времени можно использовать флуоресцентный краситель, который образует комплекс краситель-ДНК (например, SYBR® Green) или помеченный флуорофором ДНК-зонд, такой как одноцепочечные олигонуклеотиды, ковалентно связанные с флуоресцентным репортером или флуорофором (например, 6-карбоксифлуоресцеин или тетрафтор-флуоресцеин), и тушитель флуоресценции (например, тетраметилродамин или дигидроциклопирролоиндол-трипептид – молекулы, связывающиеся с малой бороздкой ДНК). Флуоресцентный сигнал позволяет определить амплифицированный продукт в режиме реального времени, при этом указывая на присутствие представляющей интерес последовательности и позволяя провести количественный анализ числа копий рассматриваемой последовательности в клеточной ДНК или уровень экспрессии рассматриваемой последовательности от клеточной мРНК.

Идентификация полиморфизмов длины рестрикционных фрагментов (ПДРФ) рассматривается, например, в работах Alonso-Blanco et al. (*Methods in Molecular Biology*, vol.82, "Arabidopsis Protocols", pp. 137-146, J.M. Martinez-Zapater and J. Salinas, eds., c. 1998 by Humana Press, Totowa, NJ); Burr ("Mapping Genes with Recombinant Inbreds", pp. 249-254, in Freeling, M. and V. Walbot (Ed.), *The Maize Handbook*, c. 1994 by Springer-Verlag New York, Inc.: New York, NY, USA; Berlin Germany; Burr et al. *Genetics* (1998) 118: 519 и Gardiner, J. et al., (1993) *Genetics* 134: 917). Например, чтобы создать библиотеку для ПДРФ, расширенную однокопийными или малокопийными экспрессированными последовательностями, всю ДНК можно расщепить чувствительными к метилированию ферментами (например, *PstI*). Расщепленную ДНК можно разделить по размеру на препаративном геле. Фрагменты полинуклеотидной цепи (от 500 до 2000 п.о.) могут быть отделены, извлечены и клонированы в плазмидный вектор (например, pUC18). Саузерн-блоттинг плазмидных дайджестов можно зондировать со всей дегидратированной в результате гидродинамического сдвига ДНК для

отбора клонов, которые гибридизируются в однокопийные или малокопийные последовательности. Для увеличения числа обнаруженных полиморфизмов можно провести тестирование дополнительных рестрикционных эндонуклеаз.

Идентификация ПДАФ рассматривается, например, в Европейском патенте 0 534 858 и в патенте США № 5,878,215. В целом, вся клеточная ДНК расщепляется с помощью одного или более рестрикционных ферментов. Полусайт-специфические адаптеры рестрикции лигируются во все фрагменты рестрикции, и эти фрагменты селективно амплифицируются двумя праймерами для ПЦР, которые имеют соответствующий адаптер и последовательности сайт-специфической рестрикции. Продукты ПЦР можно визуализировать после фракционирования по размеру, как описано выше.

В некоторых вариантах методы направлены на селекцию линии растений. В таких методах используются генетические полиморфизмы, определяемые описанным выше способом в программе скрещивания с использованием маркера, для облегчения выведения линий, которые имеют желаемые изменения в характеристике биомассы. При определении подходящего генетического полиморфизма, связанного с вариацией признаков, определяется одно или несколько отдельных растений, которые обладают полиморфной аллелью, коррелирующей с желаемой вариацией. Затем такие растения используют в программе разведения, чтобы объединить полиморфную аллель с многочисленными другими аллелями в других локусах, которые коррелируют с желаемой вариацией. Способы, пригодные для использования в программах селекции растений, известны в данной отрасли знаний и включают в том числе обратное скрещивание, массовую селекцию, селекционное разведение, смешанную селекцию, скрещивание с другой популяцией и периодическую селекцию. В селекционной программе данные способы можно использовать по отдельности или в сочетании с одним или несколькими другими способами. Таким образом, происходит самовоспроизводство каждого определенного растения или его скрещивание с другим растением для получения семян, которые затем проращиваются и формируют потомственные растения. Затем по меньшей мере одно из таких потомственных растений

самовоспроизводится и скрещивается с другим растением для формирования последующего поколения потомков. В программе разведения при необходимости можно повторять шаги самовоспроизводства или ауткроссинга для получения дополнительно от 0 до 5 поколений, чтобы достичь желаемого единообразия и стабильности получаемой в результате линии растений, которая сохраняет полиморфную аллель. В большинстве программ разведения анализ конкретной полиморфной аллели проводится в каждом поколении, хотя при желании анализ можно проводить в чередующихся поколениях.

В некоторых случаях также проводится селекция по другим полезным признакам, например селекция по устойчивости к грибкам или бактериям. Селекцию по таким другим признакам можно проводить до, во время или после идентификации отдельных растений, которые обладают требуемой полиморфной аллелью.

VII. Изделие

Трансгенные растения, приведенные в настоящем документе, находят различное применение в сельском хозяйстве и в отраслях, связанных с выработкой энергии. Например, трансгенные растения, описанные в настоящем документе, можно использовать для производства кормов для животных и пищевых продуктов. Такие растения при этом часто особенно полезны в качестве исходного сырья для производства энергии.

Трансгенные растения, описанные в настоящем документе, часто дают более высокий урожай зерна и/или выход биомассы на гектар по сравнению с контрольными растениями, в которых отсутствует экзогенная нуклеиновая кислота. В некоторых вариантах такие трансгенные растения обеспечивают эквивалентный или даже более высокий урожай зерна и/или выход биомассы на гектар по сравнению с контрольными растениями при выращивании в условиях уменьшенного потребления, например, удобрений и/или воды. Таким образом, такие трансгенные растения можно использовать для обеспечения стабильного урожая при более низкой производственной себестоимости и/или в тяжелых условиях окружающей среды, таких как засуха. В некоторых вариантах состав растений, описанных в настоящем документе, обеспечивает их более

эффективную переработку в свободные сахара и затем в этанол с целью производства энергии. В некоторых вариантах такие растения обеспечивают более высокий выход этанола, бутанола, диметилового эфира, других молекул биотоплива и/или полученных из сахара побочных продуктов на килограмм растительного материала по сравнению с контрольными растениями. При такой эффективной переработке из растения извлекаются в том числе материалы, содержащие глюкан, целлюлозу, гемицеллюлозу и лигнин. Обеспечивая более высокий выход по биомассе при эквивалентных или даже уменьшенных издержках производства, трансгенные растения, описанные в настоящем документе, повышают рентабельность фермерских хозяйств и перерабатывающих предприятий, а также снижают затраты потребителей.

Семена трансгенных растений, описанных в настоящем документе, можно приводить в определенное состояние и упаковывать в пакеты способами, известными в данной отрасли, для формирования изделия. Упаковочный материал, такой как бумага и ткань, хорошо известны в данной отрасли. На пакете с семенами может быть маркировка, например этикетка или бирка, прикрепленная к упаковочному материалу, этикетка, напечатанная на упаковочном материале, или вложенная в пакет этикетка с описанием происхождения семян, находящихся в нем.

Далее изобретение будет далее описано на следующих примерах, которые не ограничивают объем изобретения, описанного в патентной формуле.

VIII. Примеры

Пример 1. Трансгенные растения риса

Применительно к трансформации риса используются следующие условные обозначения: T_0 : растение, регенерированное из трансформированной культуры ткани; T_1 : потомство первого поколения самоопыленных растений T_0 ; T_2 : потомство второго поколения самоопыленных растений T_1 ; T_3 : потомство третьего поколения самоопыленных растений T_2 .

Ниже приведен перечень нуклеиновых кислот, выделенных из растений *Arabidopsis thaliana*: CeresAnnot: 544549 (SEQ ID NO: 262), CeresAnnot: 1355066 (SEQ ID NO: 116), CeresClone: 1356785

(SEQ ID NO: 252), CeresClone: 26006 (SEQ ID NO: 594), CeresClone: 4831 (SEQ ID NO: 76), CeresAnnot: 847799 (SEQ ID NO: 208) и CeresAnnot: 878355 (SEQ ID NO: 425). Приведенные ниже нуклеиновые кислоты выделены из растений *Zea mays*: CeresClone: 1384304 (SEQ ID NO: 553). Приведенные ниже нуклеиновые кислоты выделены из растений *Oryza sativa*: антисмысловая последовательность (SEQ ID NO: 678). CeresClone: 638126 (SEQ ID NO: 322) выделена из растений *Glycine max*.

Каждая изолированная нуклеиновая кислота, описанная выше, была клонирована в плазмидный вектор Ti, содержащий ген фосфинотрицин-ацетилтрансферазы, который передает трансформированным растениям устойчивость к гербициду Finale™. Конструкции составлялись с использованием вышеупомянутых нуклеиновых кислот, каждая из которых содержала в себе функционально связанный с ней промотор 326, введенный в клетки каллуса риса сорта Kitaake по протоколу агробактериальной трансформации. В результате каждой трансформации было получено приблизительно 20-30 независимых трансгенных растений T₀, в том числе для контрольной плазмиды (пустого вектора). Предварительный анализ фенотипа продемонстрировал, что трансформанты T₀ не показали каких-либо существенных фенотипических аномалий в вегетативных органах, за небольшими исключениями, когда некоторые растения оказались маленькими, с пониженной фертильностью, наиболее вероятно, вследствие воздействия культуры тканей.

Растения T₀ выращивались в теплице, что обеспечивало самоопыление и сбор семян T₁. Растения T₁ и T₂ выращивались в поле. Наличие каждой конструкции подтверждалось ПЦР.

Перед весенним прорастанием семена риса вымачивались в течение 3-4 дней и приблизительно через один месяц пересаживались в поле в г. Лангфан, Китай. Расстояние между рядами составляло 25 см, расстояние между растениями составляло 15 см. Непосредственно перед посадкой было внесено комплексное удобрение (16N-16P-16K) из расчета 25 кг/му (666,7 м²). В период вегетации до развития метелки в два приема вносилась мочеви́на из расчета 12,5 кг/му.

В одном ряду на каждый трансгенный объект росло десять растений. Измерения выполнялись только в тех рядах, где визуально было заметно отличие от контрольных растений. Высота растений измерялась по достижении зрелости.

Данные измерений биомассы (сухая масса) CW00733, CW00710, CW00628, CW00604, CW00564, CW00469 и CW00536 снимались с выращенных растений T_1 . Стебли с листьями и листовыми влагалищами, но без метелок, высушивались в теплице не менее месяца, и затем производилось взвешивание каждого растения (все побеги каждого растения взвешивались вместе). Данные по измерениям CW00191, CW00297 и CW00319 снимались с выращенных растений T_2 .

Стебли с листьями и листовыми влагалищами, но с отделенными метелками высушивались в помещении не менее месяца, и затем производилось взвешивание каждого растения (все побеги каждого растения взвешивались вместе). Количество побегов подсчитывалось после 4 месяцев роста.

Пример 2. Результаты по объектам риса CW00733, Ceres Clone:1384304, (SEQ ID NO: 553)

Анализ семени T_1 от одного объекта CW00733, содержащего CeresClone: 1384304, проводился, как описано в примере 1. Высота растения, биомасса и масса метелки трансгенных растений T_1 в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 1. Данные каждой строки таблицы соответствуют ряду в поле. Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы, высоты и массы метелки показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 1

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
72,41	80,30	16,70	27,45	21,56	28,89
				24,58	23,42

Высота растений (см), урожайность (измеряется в г/метелки

16 растений) и биомасса (измеряется в г только стебель (без соцветия или корня)) трансгенных растений T₂ CW00733 в сравнении с растениями, не содержащими трансген и выращенными (МАССА) в том же месте, показаны в таблице 2. Результаты по объектам CW00604 (пример 7) также показаны в таблице 2. Наблюдалось увеличение высоты и биомассы.

ТАБЛИЦА 2

Растения T2 от объектов CW00733 и CW00604

	Отчет	МАССА	CW00604	CW00733
Высота, см	Отчет I	80,3	81,3	85,0
	Отчет II	73,3	79,7	84,5
	Отчет III	77,2	79,1	82,9
	Средний	76,9	80,0	84,1
Урожайность	Отчет I	395,4	357,0	356,8
	Отчет II	385,0	324,0	348,5
	Отчет III	361,6	309,0	345,3
	Средний	380,7	330,0	350,2
Биомасса	Отчет I	16,6	24,8	19,3
	Отчет II	20,0	20,8	17,7
	Отчет III	16,6	19,8	19,7
	Средний	17,7	21,8	18,9

Пример 3. Результаты по объектам риса CW00319, Ceres Annot:544549 (SEQ ID NO: 262)

Анализ биомассы растений, выращенных из семян T₂ и T₃ от одного объекта CW00319, содержащего Ceres Annot: 544549, проводился, как описано в примере 1. Среднее количество биомассы трансгенных растений T₂ и T₃ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показано в таблице 3. Делянки с низким содержанием азота и контрольные делянки повторялись 3 раза по схеме рандомизированных блоков, при этом были представлены трансгенные растения в виде нескольких объектов и контрольные группы. Каждая делянка содержала 40 растений. По одному объекту CW00319 проводилось измерение десяти растений на каждой делянке. Каждое значение биомассы в таблице 3 представляет собой среднюю величину по 30

измеренным растениям. Результаты показывают увеличение измеренной биомассы у трансгенных растений, выращенных в условиях низкого содержания азота в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 3

Биомасса, г/растение (нормальное)		Биомасса, г/растение (низкое содержание азота)	
Контрольное	Трансгенное	Контрольное	Трансгенное
14,6	14,38	16,01	17,44

Пример 4. Результаты по объектам риса CW00710, Ceres Annot: 1355066, (SEQ ID NO: 116)

Анализ семени T₁ от одного объекта CW00710, содержащего Ceres Annot: 1355066, проводился, как описано в примере 1. Высота растений, биомасса и масса метелки трансгенных растений T₁ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 4. Данные в строках таблицы соответствуют ряду в поле. Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы и высоты показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 4

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
72,20	85,60	14,51	24,15	19,38	18,70

Пример 5. Результаты по объектам риса CW00628 (SEQ ID NO: 678)

Анализ семени T₁ от одного объекта CW00628, содержащего РНК-интерференцию с SEQ ID NO: 678, проводился, как описано в примере 1. Высота растений, биомасса и масса метелки трансгенных растений T₁ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 5. Данные в строках таблицы соответствуют ряду в поле. Данные результатов обработки представляют средние показатели по 10 трансгенным

растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы и высоты показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 5

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
75,30	81,60	15,92	29,96	24,53	20,38

Пример 6. Результаты по объектам риса CW00297, Ceres Clone:625057, (SEQ ID NO: 644)

Анализ биомассы растений, выращенных из семян T₂ и T₃ от одного объекта CW00297, содержащего Ceres Clone: 625057, проводился, как описано в примере 1. Среднее количество биомассы трансгенных растений T₂ и T₃ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 6. Делянки с низким содержанием азота и контрольные делянки повторялись 3 раза по схеме рандомизированных блоков, при этом были представлены трансгенные растения в виде нескольких объектов и контрольные группы. Каждая делянка содержала 40 растений. По одному объекту CW00297 проводилось измерение десяти растений на каждой делянке. Каждое значение биомассы в таблице 6 представляет собой среднюю величину по 30 измеренным растениям. Результаты показывают увеличение измеренной биомассы у трансгенных растений, выращенных в условиях нормального и низкого содержания азоте в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 6

Биомасса, г/растение (нормальное)		Биомасса, г/растение (низкое содержание азота)	
Контрольное	Трансгенное	Контрольное	Трансгенное
14,6	15,93	16,01	17,35

Пример 7. Результаты по объектам риса CW00604, Ceres Clone:1356785, (SEQ ID NO: 252)

Анализ семени T₁ от одного объекта CW00604, содержащего Ceres Clone: 1356785, проводился, как описано в примере 1.

Высота растений, биомасса и масса метелки трансгенных растений T_1 в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 7. Данные каждой строки таблицы соответствуют ряду в поле. Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы, высоты и массы метелки показано в сравнении с растениями, не содержащими трансген. Также наблюдалось увеличение высоты и биомассы у растений T_2 . См. таблицу 2.

ТАБЛИЦА 7

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
72,53	81,25	16,40	28,68	22,21	29,91
73,70	82,30				

Пример 8. Результаты по объектам риса CW00564, Ceres Clone: 638126 (SEQ ID NO: 322)

Анализ семени T_1 от одного объекта CW00564, содержащего Ceres Clone: 638126, проводился, как описано в примере 1. Высота растений, биомасса и масса метелки трансгенных растений T_1 в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 8. Данные в строках таблицы соответствуют ряду в поле. Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы, высоты и массы метелки показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 8

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
72,88	85,44	18,11	36,47	22,56	37,47

Высота растения (см), урожайность (измеряется в г/метелки 16 растений) и биомасса (измеряется в г только стебель (без соцветия или корня)) трансгенных растений T_2 CW00564 в сравнении

с растениями, не содержащими трансген и выращенными (МАССА) в том же месте, показаны в таблице 9. Результаты по объектам CW00469 (пример 10) также показаны в таблице 9. Наблюдалось увеличение высоты, урожайности и биомассы.

ТАБЛИЦА 9

Растения T2 от объектов CW00564 и CW00469

	Отчет	МАССА	CW00469	CW00564
Высота, см	Отчет I	77,3	117,1	90,7
	Отчет II	76,9	117,7	91,5
	Отчет III	77,4	113,1	92,7
	Средний	77,2	115,9	91,6
Урожайность	Отчет I	356,9	426,0	406,8
	Отчет II	343,2	378,8	423,6
	Отчет III	366,7	398,0	–
	Средний	355,6	400,9	415,2
Биомасса	Отчет I	17,1	29,7	25,2
	Отчет II	15,1	26,5	25,1
	Отчет III	14,6	27,1	24,4
	Средний	15,6	27,7	24,9

Пример 9. Результаты по объектам риса CW00010, Ceres Clone: 26006 (SEQ ID NO: 594)

Анализ семени T₁ от трех объектов CWCW00010, содержащих Ceres Clone: 26006, проводился, как описано в примере 1. Высота растений, биомасса, количество побегов, период цветения и масса метелки трансгенных растений T₁ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблицах 10, 11 и 12. Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы, высоты, количества побегов и массы метелки показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 10

Объект 1

	Процентное увеличение	Значение p	Количество измеренных растений
Биомасса	9	0,307	11
Высота растения	7	0,028	11
Количество побегов	27	0,002	10
Период цветения	10	0,048	11
Масса метелки	16	0,012	39

ТАБЛИЦА 11

Объект 17

	Процентное увеличение	Значение p	Количество измеренных растений
Высота растения	3	0,004	7
Количество побегов	34		1
Масса метелки	4	0,591	7

ТАБЛИЦА 12

Объект 2

	Процентное увеличение	Значение p	Количество измеренных растений
Высота растения	2	0,007	15
Количество побегов	27	0,013	10
Масса метелки	16	0,011	14

Пример 10. Результаты по объектам риса CW00469, Ceres Clone: 4831 (SEQ ID NO: 76)

Анализ семени T_1 от одного объекта CW00469, содержащего Ceres Clone: 4831, проводился, как описано в примере 1. Высота растений, биомасса и масса метелки трансгенных растений T_1 в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 13. Данные в строках таблицы соответствуют ряду в поле.

Данные результатов обработки представляют средние показатели по 10 трансгенным растениям (1 ряд одного объекта) и средние показатели 40 контрольных растений (4 ряда). Увеличение биомассы, высоты и массы метелки показано в сравнении с растениями, не содержащими трансген. Увеличение высоты, урожайности и биомассы показано у растений T_2 (см. таблицу 9).

ТАБЛИЦА 13

Высота растения, см		Биомасса, г/растение		Масса метелки, г/растение	
Контрольное	Трансгенное	Контрольное	Трансгенное	Контрольное	Трансгенное
76,54	112,22	25,44	56,06	32,63	62,70

Пример 11. Результаты по объектам риса CW00536, Ceres Annot: 847799, (SEQ ID NO: 208)

Анализ семени T_1 от 16 объектов CW00536, содержащего Ceres Annot: 847799, проводился, как описано в примере 1. Высота растений и масса метелки трансгенных растений T_1 в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показана в таблицах 14 и 15. Данные результатов обработки представляют средние показатели по 16 объектам с 15 трансгенным растениями на каждый объект и средние показатели нескольких сотен контрольных растений. Увеличение высоты и массы метелки показано в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 14

Масса метелки, г/растение			
Контрольное	Коэффициент Стьюдента	Трансгенное	Коэффициент Стьюдента
17,916	2,181	20,854	3,419

ТАБЛИЦА 15

Высота растения, см	
Объект	Процентное увеличение по сравнению с контрольным
CW00536-03	7,79
CW00536-05	5,66
CW00536-11	8,71
CW00536-12	8,47
CW00536-20	8,77

Пример 12. Результаты по объектам риса SR05004, CW00191, CeresAnnot: 878355 (SEQ ID NO: 425)

Анализ биомассы растений, выращенных из семян T₂ и T₃ от одного объекта CW00191, содержащего Ceres Annot: 878355, проводился, как описано в примере 1. Среднее количество биомассы трансгенных растений T₂ и T₃ в сравнении с растениями, не содержащими трансген и выращенными в том же месте, показаны в таблице 16. Делянки с низким содержанием азота и контрольные делянки повторялись 3 раза по схеме рандомизированных блоков, при этом были представлены трансгенные растения в виде нескольких объектов и контрольные группы. Каждая делянка содержала 40 растений. По одному объекту CW00191 проводилось измерение десяти растений в каждой делянке. Каждое значение биомассы в таблице 16 представляет собой среднюю величину по 30 измеренным растениям. Результаты показывают увеличение измеренной биомассы у трансгенных растений, выращенных в условиях нормального и низкого содержания азота в сравнении с растениями, не содержащими трансген.

ТАБЛИЦА 16

Биомасса, г/растение (нормальное)		Биомасса, г/растение (низкое содержание азота)	
Контрольное	Трансгенное	Контрольное	Трансгенное
14,6	16,75	16,01	19,64

Пример 13. Определение функциональных гомологов с помощью программы Reciprocal BLAST

Отбираемая последовательность рассматривалась как функциональный гомолог эталонной последовательности, если отбираемая и эталонная последовательности, которые кодировали белки, выполняли аналогичную функцию и/или действие. Для определения потенциальных последовательностей функциональных гомологов в базах данных всех общедоступных и запатентованных пептидных последовательностей, в том числе всего банка данных Национального центра биотехнологической информации, а также синтеза пептидов из клонов Ceres, использовался так называемый процесс Reciprocal BLAST (двусторонний процесс поиска с помощью программы BLAST) (Rivera et al., *Proc. Natl. Acad. Sci. USA*, 95:6239-6244 (1998)).

Перед запуском процесса Reciprocal BLAST был произведен поиск специфического эталонного полипептида по всем полипептидам из исходных видов с использованием средства поиска основного локального выравнивания (BLAST) с целью определения полипептидов, имеющих идентичность последовательности согласно программе BLAST 80% или более по отношению к эталонному полипептиду и длину выравнивания 85% или больше вдоль более короткой последовательности в линии. Эталонный полипептид и прочие упомянутые выше идентифицированные полипептиды были обозначены как кластер.

Для определения в банке данных идентичности последовательностей и значения E была использована программа BLASTP версии 2.0 Вашингтонского университета, Сент-Луис, штат Миссури, США. Параметры программы BLASTP версии 2.0 следующие: 1) предельное значение E: 1.0e-5; 2) разрядность: 5; 3) опция - postsw. Идентичность последовательностей согласно программе

BLAST вычислялась, исходя из выравнивания первых BLAST HSP (пар сегментов с максимальным сходством), идентифицированных потенциальных последовательностей функциональных гомологов с конкретным эталонным полипептидом. Число однозначно совпавших остатков при выравнивании BLAST HSP (пар сегментов с максимальным сходством) делилось на длину HSP и затем умножалось на 100 для получения идентичности последовательностей BLAST. Длина пар сегментов с максимальным сходством (HSP) обычно включала разрывы в последовательности, однако в некоторых случаях разрывы исключались.

Основной процесс Reciprocal BLAST состоит из двух этапов: поиска в банке данных с помощью программы BLAST в прямом направлении и поиска в обратном направлении. Во время поиска в прямом направлении производилось сравнение эталонной последовательности полипептида – «полипептида А» из исходных видов SA со всеми белковыми последовательностями из представляющего интерес вида. Максимальные совпадения определялись с использованием предельного значения E 10^{-5} и предельного значения идентичности последовательности 35%. Среди максимальных совпадений последовательность с наименьшим значением E обозначалась как лучшее совпадение и рассматривалась как потенциальный функциональный гомолог или ортолог. Любое другое максимальное совпадение, при котором идентичность последовательности была 80% или более по отношению к лучшей последовательности или к эталонному полипептиду, также рассматривалось как потенциальный функциональный гомолог или ортолог. Данный процесс повторялся для всех представляющих интерес видов.

Во время поиска в обратном направлении производилось сравнение максимальных совпадений, обнаруженных при поиске в прямом направлении, из всех видов, со всеми белковыми последовательностями из исходных видов SA. Максимальное совпадение из результатов поиска в прямом направлении, которое возвращало полипептид из вышеупомянутого кластера как лучшее совпадение, также рассматривалось как потенциальный функциональный гомолог.

Функциональные гомологи идентифицировались путем проверки вручную последовательностей потенциальных функциональных гомологов. Репрезентативные функциональные гомологи последовательностей SEQ ID NO: 554, 263, 117, 1, 645, 253, 323, 595, 77, 209 и 426 показаны, соответственно, на фиг. 1-11. Дополнительные примеры гомологов взаимосвязаны с определенными фигурами в перечне последовательностей.

Пример 14. Определение функциональных гомологов с помощью скрытой марковской модели

Скрытые марковские модели (НММ) создавались программой HMMER 2.3.2. Для создания каждой модели НММ использовались параметры программы HMMER 2.3.2 по умолчанию, сконфигурированные для глобального выравнивания.

Модель НММ создавалась с использованием в качестве входных данных последовательностей, показанных на фиг. 1. Данные последовательности согласовывались с моделью, а репрезентативный показатель НММ для каждой последовательности показан в перечне последовательностей. Дополнительные последовательности согласовывались с моделью, а репрезентативные показатели НММ для каких-либо дополнительных последовательностей показаны в перечне последовательностей. Результаты показали, что данные дополнительные последовательности являются функциональными гомологами последовательности SEQ ID NO: 554.

Указанная выше процедура была повторена, и была создана модель НММ для каждой группы последовательностей, показанных на фиг. 2-11, с использованием последовательностей, показанных на каждом чертеже в качестве входных данных для этой модели НММ. Репрезентативный показатель по каждой последовательности показан в перечне последовательностей. Дополнительные последовательности согласовывались с определенными моделями НММ, а репрезентативные показатели НММ для таких дополнительных последовательностей показаны в перечне последовательностей. Результаты показывают, что данные дополнительные последовательности являются функциональными гомологами последовательностей, использованных для создания данной модели НММ.

Другие варианты

Следует понимать, что хотя описание изобретения дано в сочетании с его полным описанием, вышеизложенное описание предназначено для иллюстрации и не ограничивает объем изобретения, который определяется объемом прилагаемой формулы изобретения. Прочие аспекты, преимущества и модификации находятся в пределах следующей формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ получения растения с повышенным уровнем биомассы, где способ включает выращивание растения из растительной клетки, содержащей экзогенную нуклеиновую кислоту, содержащую регуляторную область, функционально связанную с нуклеотидной последовательностью, кодирующей полипептид, модулирующий уровень биомассы растения, последовательность которого на 85% или более идентична аминокислотной последовательности SEQ ID NO:77, и отбор растения с повышенным уровнем биомассы по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанной экзогенной нуклеиновой кислоты.

2. Способ по п.1, где последовательность указанного полипептида на 90% или более идентична остаткам 57-127 последовательности SEQ ID NO:77.

3. Способ получения растения с повышенным уровнем биомассы, включающий выращивание растения из растительной клетки, содержащей экзогенную нуклеиновую кислоту, которая содержит регуляторную область, функционально связанную с нуклеотидной последовательностью, которая на 80% или более идентична нуклеотидной последовательности SEQ ID NO:76, которая кодирует полипептид, модулирующий уровень биомассы, и отбор растения с повышенным уровнем биомассы по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанной экзогенной нуклеиновой кислоты.

4. Способ повышения уровня биомассы у растения, включающий введение в растительную клетку экзогенной нуклеиновой кислоты, содержащей регуляторную область, функционально связанную с нуклеотидной последовательностью, кодирующей полипептид, модулирующий уровень биомассы растения, последовательность которого на 85% или более идентична аминокислотной последовательности SEQ ID NO:77, и где отбирают растение, выращиваемое из указанной растительной клетки, с повышенным уровнем биомассы по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанной нуклеиновой кислоты.

5. Способ по любому из пп.1, 3 или 4, где последовательность указанного полипептида представляет собой SEQ ID NO:77.

6. Способ повышения уровня биомассы растения, включающий введение в растительную клетку экзогенной нуклеиновой кислоты, содержащей регуляторную область, функционально связанную с нуклеотидной последовательностью, которая на 80% или более идентична нуклеотидной последовательности SEQ ID NO:76, которая кодирует полипептид, модулирующий уровень биомассы растения, и где отбирают растение, выращиваемое из указанной растительной клетки, с повышенным уровнем биомассы по сравнению с соответствующим уровнем контрольного растения, которое не содержит указанной нуклеиновой кислоты.

7. Растительная клетка, включающая экзогенную нуклеиновую кислоту, содержащую регуляторную область, функционально связанную с нуклеотидной последовательностью, кодирующей полипептид, модулирующий уровень биомассы растения, последовательность которого на 85% или более идентична аминокислотной последовательности SEQ ID NO:77, где растение, выращиваемое из указанной растительной клетки, имеет повышенный уровень биомассы растения по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанной нуклеиновой кислоты.

8. Растительная клетка, включающая экзогенную нуклеиновую кислоту, содержащую регуляторную область, функционально связанную с нуклеотидной последовательностью, которая на 80% или более идентична нуклеотидной последовательности SEQ ID NO:76, которая кодирует полипептид, модулирующий уровень биомассы растения, и где растение, выращиваемое из указанной растительной клетки, имеет повышенный уровень биомассы по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанной нуклеиновой кислоты.

9. Трансгенное растение, содержащее растительную клетку по п.7 или 8, где указанное растение имеет повышенный уровень биомассы по сравнению с соответствующим уровнем биомассы контрольного растения, которое не содержит указанную нуклеиновую

кислоту.

10. Трансгенное растение по п.9, выбранное из группы, включающей *Panicum virgatum* (просо), *Sorghum bicolor* (сорго, суданская трава), *Miscanthus giganteus* (мискантус), *Saccharum* sp. (сахарный тростник), *Populus balsamifera* (тополь бальзамический), *Zea mays* (маис), *Glycine max* (соя), *Brassica napus* (канола), *Triticum aestivum* (пшеница), *Gossypium hirsutum* (хлопчатник обыкновенный), *Oryza sativa* (рис), *Helianthus annuus* (подсолнечник однолетний), *Medicago sativa* (люцерна), *Beta vulgaris* (сахарная свекла) или *Pennisetum glaucum* (просо африканское).

11. Трансгенное растение по п.9, содержащее полипептид, модулирующий уровень биомассы растения, имеющий аминокислотную последовательность SEQ ID NO:77.

12. Семя трансгенного растения по п.11.

13. Семя, которое дает трансгенное растение по п.9.

По доверенности

SEQ_ID_NO_567	MVRT	-----	-----	-	PERSY	-	SC	SSAKEVAYSC	GYCGYALNLS	31
SEQ_ID_NO_569	MARA	-----	-----	-	LDRE	-----	Y	TSGKDVAYSC	GYCGYALNLS	28
SEQ_ID_NO_554	MDRSASVGI	K	DGGFGGNHLY		SPSFS	-----	SS	SSMRHVNYS	GSCGYELNLS	47
SEQ_ID_NO_556	MERSASVGV		DGRFGGNQFY		SPSFS	SSSSS		SSMRHVNYS	GSCGYELNLS	50
SEQ_ID_NO_563	MENS	M-VLAH	E	-----	TQ	-----		SSRDVSYSC	GSCGYELNLS	39
SEQ_ID_NO_560	MDDSA	EKR	G	-----	GH	-----		SSQRDVCS	GTGCGYELNLS	38
SEQ_ID_NO_565	MDLSH	DLAK	V	-----	-----	-----	DS	MPNKNSSPC	GSCGYELNLS	32
SEQ_ID_NO_558	MEKSK	VGKG	A-YLNGN	-----	-----	-----		ASQRHVSYS	GSCGYELNLS	44
SEQ_ID_NO_562	MEKSV	FVKD	MOQLNGN	-----	-----	-----	SS	VSQRDVTYS	GSCGYELNLS	43

SEQ_ID_NO_567	SSTRNTANI	G	SKYGKQ	RKG	VI	SFFAI	DES	RFTQTDEVSC	MPYFHSRRSW	81
SEQ_ID_NO_569	SSARNTANI	G	SKYGKH	RKG	VVSFFAI	DES		RFTQTDEVSC	TPYFHS	78
SEQ_ID_NO_554	STNRITSSI	G	SKYGKSMK	TG	II	SFFNI	DEG	RFSQVDEFQC	MPHFSRYSW	96
SEQ_ID_NO_556	STNRITSTI	G	SKYGKSMK	SG	II	SFFNI	DEG	RFSQVDEFQC	MPHFSRYSW	99
SEQ_ID_NO_563	SSNRNT	KNI	G	SKYGKNI	KRG	VI	SFLQVDES	RFTLVDEFKC	IPYFHSRHSW	89
SEQ_ID_NO_560	SSNRNTASI	G	SKYGKSI	KRG	II	SFFSI	DS	RFTQVDEFQC	VPHFDSKHSW	87
SEQ_ID_NO_565	SSSRNTASI	G	SKYGKSI	KKG	MI	SFLSI	DES	RFTQVDEFKC	VPFFFSKRWS	82
SEQ_ID_NO_558	SSNRNTSSI	G	SKYGKSI	KRG	II	SFFFI	DES	RFTQVDEFQC	LPFFSRNSW	93
SEQ_ID_NO_562	SSSRNTATI	G	SKYGKI	KRG	MI	SFFNI	DET	RFTQVDEFQC	RPYFSKHSW	92

SEQ_ID_NO_567	GLFRK	RTLRI	CRKCGGR	IGN	A	YEDED	---	STLY	DGSDDL	HMSSE	GYS	M	126
SEQ_ID_NO_569	GFR	RNRTRL	CRKCSGHI	GN	A	YEDED	---	PTLC	DGSDDL	DMSSK	GSS	T	123
SEQ_ID_NO_554	GLFRRKTKLL		CRQCNNYI	GN	AS	YDKAPPEY		ALVTQNS	S---	---	PKGV	T	140
SEQ_ID_NO_556	GLFRHRTKLL		CRKCNNYI	GN	A	SQEKAPPEY		ALVTQNSD	---	---	PTSPR	I	142
SEQ_ID_NO_563	GLFCHRTKLL		CRKCSNHVGN		A	YVRKSP	Y	PLVLDES	SDSA	PAPATAT	TDI		137
SEQ_ID_NO_560	GLFRRRTKLL		CRKCGNHI	GN	A	YNGYTSS	F	PLVSDGAE	---	---	SSPSSK	V	131
SEQ_ID_NO_565	GLFQRRTKLQ		CRKCGND	GI	A	YDDSA	SSY	PLVADASD	---	---	TASGSE	I	126
SEQ_ID_NO_558	GLFHRRRTALL		CRKCGNNI	GI	A	YDDKASAY		PLVADGSD	---	---	SSSVSE	V	137
SEQ_ID_NO_562	GLFRHRTKLL		CRKCGNHI	GD	A	YDDKSSGY		PHVLDGSD	---	---	SSSGTE	P	136

SEQ_ID_NO_567	SSGKKYVI	KI	NALQPSTDDS		-	GVPFTL	152
SEQ_ID_NO_569	SPRKKYVI	KI	NALQPSSDDS		-	GALFSP	149
SEQ_ID_NO_554	DTVTKYDI	RI	RALQPSS	---	-	GVASL	162
SEQ_ID_NO_556	GSVTKYDI	RI	RSLQPSS	---	-	AVALL	164
SEQ_ID_NO_563	STCRKYDI	RL	RALQPSSSEP		SAI	PLLT	164
SEQ_ID_NO_560	VSHTKYDI	RI	CALQPSSSEE		SGI	PVFA	158
SEQ_ID_NO_565	TTHRKYNI	KI	RSLQPSSSAS		-	GTPLPE	152
SEQ_ID_NO_558	SKHRKYDVKI		RALQPSSVDQ		FSTPI	HT	164
SEQ_ID_NO_562	SNHRKYDVRI		RALQPSTAEG		LGS	PLFA	163

ФИГ. 1

SEQ_ID_NO_263	-----	-----	-----	-----	---MMLSHLF	7
SEQ_ID_NO_269	-----	-----	-----	-----	-----	0
SEQ_ID_NO_264	-----	-----	-----	-----	-----	0
SEQ_ID_NO_266	-----	-----	-----	-----	-----	0
SEQ_ID_NO_268	-----	-----	-----	-----	-----	0
SEQ_ID_NO_271	MDAGTARQLE	TIKVMSVVAV	SSVPLCLEKL	SAYLDEKRVA	VGKLGV DGGQ	50
SEQ_ID_NO_273	-----	-----	-----	-----	-----	0

SEQ_ID_NO_263	MSLSKLT CNF	SIFSVFMACG	SLGMSQVRDT	PVKLFGMTIT	PVSHDP - - -	53
SEQ_ID_NO_269	-----	-----	-----MDDL	AAASPPHPP	PPPPES - - -	19
SEQ_ID_NO_264	-----	-----	-----MLDFKDP	AIKLFGKTI	SLPLNP - - -	22
SEQ_ID_NO_266	-----	-----	-----MSDL	AIKLFGKTI	PLQVNQLEGD	23
SEQ_ID_NO_268	-----	-----	-----MGEGRAGDG	LIKLFGKTI	PVPETA - - -	24
SEQ_ID_NO_271	QQA KPGACVC	LDRSEKKMMA	ECQGGGGGDF	LIKLFGKTI	PVPESG - - -	95
SEQ_ID_NO_273	-----	-----MVV	ESAGGGGGDC	LIKLFGKTI	PVPEAA - - -	28

SEQ_ID_NO_263	-----	-----	-----	-----	YSSSSHVL PD	63
SEQ_ID_NO_269	-----	-----HVPPP	PQ - - - -	-----TPE	KDSCEDTGDM	39
SEQ_ID_NO_264	-----	-----HLSPTSPPP	PPLSSTTSFP	-----DDT	SQGLQPPSQD	54
SEQ_ID_NO_266	VSCAI NKHES	SSGTPVAPPE	PEEYCYTATT	CLHEKDKEEQ	EGNHRESSGE	73
SEQ_ID_NO_268	-----	-----	-----	-----	AVGEAAKD	32
SEQ_ID_NO_271	-----	-----	-----	-----	-----	95
SEQ_ID_NO_273	-----	-----	-----	-----	-----AD	30

SEQ_ID_NO_263	SSSSSSSSSL	SLRPHMMNNQ	SVTDNTSLKL	SSNLNNEske	TSENSDDQHS	113
SEQ_ID_NO_269	RISEEKPCTD	QELDADQMNS	SSFNSSSECE	NLQTPSNDEM	TGSESKSEAA	88
SEQ_ID_NO_264	QEFEGKEEDG	TSRQTSEELK	DPTASPGVSE	NPETPSADKE	TL - - - SKDG	99
SEQ_ID_NO_266	EIANEKQEDV	TSCQITEDSK	DPTSSGI SE	NPKTPSVERE	TSSLKSSKDE	122
SEQ_ID_NO_268	MQQSGGSGSG	TTDPKGGQEN	-----NVQD	STGSPFQQEV	ADTEDSSADK	75
SEQ_ID_NO_271	-----DA	KENPHPLDSS	DPSPQPEVVD	A-EDPKSSPE	TT HQKPGQGN	135
SEQ_ID_NO_273	KESGSSSISC	TESDTPENAS	EPSPQPEVMD	A-EDPKSSPE	TL - - - QPPP	74

ФИГ. 2 (1 из 4)

SEQ_ID_NO_263	E I T T T S E E E	K T T E L K K P D K	I L P C P R C N S A	D T K F C Y Y N N Y	N V N Q P R H F C R	163
SEQ_ID_NO_269	Q T E G G S S E E	K V - - L K K P D K	I L P C P R C N S M	D T K F C Y Y N N Y	N I N Q P R H F C K	136
SEQ_ID_NO_264	E Q S E I S G S Q E	K T - - L K K P D K	I L P C P R C N S M	D T K F C Y Y N N Y	N V N Q P R H F C K	147
SEQ_ID_NO_266	E Q R E T S I S Q E	K T - - L K K P D K	I L P C P R C N S M	D T K F C Y Y N N Y	N V N Q P R H F C K	170
SEQ_ID_NO_268	Q Q G E A G N P K E	K - - - L K K P D K	I L P C P R C N S M	D T K F C Y Y N N Y	N I N Q P R H F C K	122
SEQ_ID_NO_271	G S G D A A S Q R E	K - - - L K K P D K	V L P C P R C N S M	D T K F C Y F N N Y	N V N Q P R H F C K	182
SEQ_ID_NO_273	G A G D L A G Q R E	K - - - L K R P D K	V L P C P R C N S M	D T K F C Y F N N Y	N V N Q P R H F C K	121

SEQ_ID_NO_263	K C Q R Y W T A G G	S M R L V P V G S G	R R K N K G W V S S	D Q - - - - - Y	L H I T S E D T D N	206
SEQ_ID_NO_269	S C Q R Y W T A G G	S M R N L P V G A G	R R K S K S S T A N	Y R - - - - - S	I T G S N L A A	179
SEQ_ID_NO_264	N C Q R Y W T A G G	T M R N V P V G A G	R R K N K N S S A S	Q Y - - - - - P	L Q T A R A S A A N	190
SEQ_ID_NO_266	K C Q R Y W T A G G	T M R N V P V G A G	R R K N K S S S A S	H Y R H L M V S E A	L R T V Q V H A M N	220
SEQ_ID_NO_268	N C Q R Y W T A G G	A M R N V P V G A G	R R K S K S A S A T	S H - - - - - F	L Q R V R A G L P V	165
SEQ_ID_NO_271	N C Q R Y W T A G G	A M R N V P V G A G	R R K N K N A V A A	S H - - - - - F	L H R V G A A C - G	224
SEQ_ID_NO_273	N C Q R Y W T A G G	A M R N V P V G A G	R R K N K H A V A A	S H - - - - - F	L H R V R A A L P A	164

SEQ_ID_NO_263	Y N - - - - -	- - - - S S S T K I	L S F E S S D S L V	- - - - -	- - - - -	224
SEQ_ID_NO_269	P A G D A P L Y Q L	S I K G D Q T A T A	V K F A P D S P L C	N S M A - S V L K I	G E Q S K N A - -	225
SEQ_ID_NO_264	G I H H P A L - -	- - - - G N N G T V	L N F G S D A - -	- - - - -	- - - - -	210
SEQ_ID_NO_266	G V H N P S F - -	- - - - G N N T T V	L A F G S D S P L C	D S V A - S V L N L	S E K T Q N S - -	259
SEQ_ID_NO_268	- - - D P L V C A A	A - - - K T N G T V	L S F G - - - - -	- S A M - S S L D L	T E Q M K Q L K E K	201
SEQ_ID_NO_271	G G G D T L K - -	- - - - T T N G T V	L S F G G H G G G C	V P P G P A C L D L	V E Q L S H H - -	264
SEQ_ID_NO_273	A A G D P L - - -	- - - - R A N A T V	L S F G G H G H G A	- P P A - A L Q D L	A E Q V T H L K E K	204

SEQ_ID_NO_263	- - - - -	- - - - - T E	R P K H Q S N E V K	- - - I N A E P V S	Q - - - - - E	245
SEQ_ID_NO_269	- - K P T S T A Q P	R N G E T Q T C P A	S G T - T S D S P R	N E P V N G A V S G	H - - - - - Q	264
SEQ_ID_NO_264	- - - - -	- - - - -	- - - - -	- - - E K G N I A G	L - - - - - E	219
SEQ_ID_NO_266	- - V R N E Y H R P	E H R I F V P C G G	A G S N G D D R S S	G S P A T A S D S S	E K G C N G N S R E	307
SEQ_ID_NO_268	L V P I - - - A G	D E R S V G S R T Q	G P S A K A E D P D	- - - R K E N V T A	D - - - - - K	236
SEQ_ID_NO_271	L A A P V I R N A G	N N P - - G P C S E	G S S N C R D D N K	T - - I N D R S C V	D - - - - - E	302
SEQ_ID_NO_273	L L V P A - R N A G	D P S P V G P C S E	G P S S T D D K A H	G G G I K E K P A V	D - - - - - R	245

ФИГ. 2 (2 из 4)

SEQ_ID_NO_263	PN- -NFQGLL	PPQASPVSP	- - - -PW	PY	QYL- - - -PP	NPS- - - - -	273
SEQ_ID_NO_269	NGI VGHSGVP	PMHPI PCFPG	PPFVYPW	SP	AMNGI - PAMA	PPVCTAPAE	312
SEQ_ID_NO_264	PVVKNFQAFS	PHV- -PCFPG	ASWSYPWN	PA	QWSSK- - -IP	PPA- - - - -	257
SEQ_ID_NO_266	AVNKDYQSFP	PQV- -PCFPG	P- - -PW	PY	QWNSA- - -LP	PP- - - - -	340
SEQ_ID_NO_268	SA- - - -RV	VQH- -PCMTN	- -GVAMN	- -	PF-SCAPPVP	ASAC- - - - -	266
SEQ_ID_NO_271	AAAAANGDDGS	VQH- -PASMN	NGGATVW	PP	PY-SC- - -AP	SPAA- - - - -	339
SEQ_ID_NO_273	PA- - - -NGA	PQL- -PASMN	- -GATVW	- -	PYGGC- - -AP	SPAA- - - - -	274

SEQ_ID_NO_263	- - - - -	- - - - -	- - - -FYHMPV	- - - -YWGCAL	- -P- -VWSTL	291
SEQ_ID_NO_269	ANSSDNGSTA	SVQWSMPPVM	PVPGYFPVLP	SS- -VMPFIS	PMPNGAWSSP	360
SEQ_ID_NO_264	- - - - -	- - - -FCPPGF	PIS- -FYPAPA	- - - -YWGCTV	- -P- -GSWNIP	285
SEQ_ID_NO_266	- - - - -	- - - -FYPPGF	PVS- -FYPAPT	- - - -YWGCTV	- -P- -SPWNVP	368
SEQ_ID_NO_268	- - - - -	- - - -YGP GSI	AIP- -FYPAAA	AAAA YWGMV	- -P- -GAWSGA	298
SEQ_ID_NO_271	- - - - -	- - - -YFSSGI	AIP- -IYPAAP	G- - -YWGMV	- -P- -GAWSLP	368
SEQ_ID_NO_273	- - - - -	- - - -YFSSGI	AIP- -IYPAAP	G- - -YWGMV	PAP- -GPWSLP	305

SEQ_ID_NO_263	- - - - -	- - - - -	- - - - -	- - - - -D	TSTCLGKRT	302
SEQ_ID_NO_269	WI- - - - -	QP- - - - -	- NCSVSASSP	TSTSTCSDNG	- SPVLGKHSR	392
SEQ_ID_NO_264	CL- - - - -	PP- - - - -	- - - - -TSSSP	IHSALATNHN	- SPTLGKHAR	313
SEQ_ID_NO_266	PC- - - - -	- - - - -	- - - - -VSPPST	SLKHCTLDSS	PTSTLGKHSR	396
SEQ_ID_NO_268	WP- - - - -	- - - -PHSGQSE	TGSSI TSASP	AASTKS- - -N	I CFTPGKHPR	334
SEQ_ID_NO_271	WPVQQPPSSQ	SQGPAAGLSS	STSPTTTSAP	SVSSSGAADS	HTLGLGKHPR	418
SEQ_ID_NO_273	WS- - - - -	- - - -VQ- - - -	- GLSLPTSAP	SLPSSGPDPD	- -LTLGKHPR	336

SEQ_ID_NO_263	DET- - - - -	- - - -SHETVKE	SKNAFERTSL	LLESQSI KNE	TSMATNNHVW	342
SEQ_ID_NO_269	DSK- - - - -	- - - -PQGDDKA	EKNLWPKTL	RI D- - - -D	PDEAAKSSI W	426
SEQ_ID_NO_264	DGE- - - - -	- - - - -	- - - - -TL	RI D- - - -D	PNEAAKSSI W	332
SEQ_ID_NO_266	DGSI LHPAYL	KEPSREGTKS	VKGVLVPKTS	RI D- - - -D	PSEAAKSSI W	440
SEQ_ID_NO_268	DRD- - - - -	- - - -EEGGAKG	NGKVWPKM	RI D- - - -D	VDEVARSSI L	368
SEQ_ID_NO_271	DRE- - - - -	- - - -EGDDGR	NAKVWAPKTI	RI D- - - -D	VDEVARSSI W	451
SEQ_ID_NO_273	EGD- - - - -	- - - -EGRSAHG	GGKVWAPKTI	RI D- - - -D	ADEVARSSI W	370

ФИГ. 2 (3 из 4)

SEQ_ID_NO_263	Y P V P M T - R E K	T Q - - - - -	- - - - - E F	S F F S N G A E - T	K S S N N R F - V P	373
SEQ_ID_NO_269	T T L G I E P G D R	S - - - - -	- - - - - M F	R S F Q S K P E - S	R E Q - - - - I S	453
SEQ_ID_NO_264	T T L G I K - N D -	- - - - - G S	N G G S - - - L L	K A F Q S K G D - E	K K R - - - - I A	362
SEQ_ID_NO_266	A T L G I M - S E K	S N - - - - S I	N G G G - - - L F	K G F Q S K N E - D	K N D - - - - M A	473
SEQ_ID_NO_268	S L I G I G - G D K	A G - - - K D G G	R G C K - - - L A	R V F E Q N E E A A	R T A T P H S A A I	409
SEQ_ID_NO_271	S L I G I K - G D K	A K Q - Q D D D A A	G G H K Q K Q L V G	M V F E P K R E - A	T K K P A A - M M	496
SEQ_ID_NO_273	S L I G I K - G D R	K Q D A A D H H G A	A G H K - - - H G	T V F E Q K R E - A	K K Q P - - - A M I	411

SEQ_ID_NO_263	ETYLNLQANP	AAMARSMNFR	ESI	396
SEQ_ID_NO_269	GAARVLQANP	AALSRSQSFQ	ETT	476
SEQ_ID_NO_264	EMSPVLQANP	AALSRSLNFH	ERA	385
SEQ_ID_NO_266	GRTSVLQANP	AALSRSLNFH	ENN	496
SEQ_ID_NO_268	SGLPELQGNP	AALSRSLTFQ	EGS	432
SEQ_ID_NO_271	TSSPLLHANP	VALTRSVAFQ	EGS	519
SEQ_ID_NO_273	ASSPLLHTNP	VALTRSVTFQ	EGS	434

ФИГ. 2 (4 из 4)

SEQ_ID_NO_132	H C C K K D P T V V	D L L P G V P Y N Q	Q F T N C C K G G A	L A S W A Q D P P N	S V A S F Q V S V G	174
SEQ_ID_NO_129	H C C K H T P S V V	D L L P G V P Y N Q	Q I A N C C R G G V	V S A Y G Q D P A G	A L S A F Q V S V G	150
SEQ_ID_NO_125	H C C K R T P A V V	D L L P G V P Y N Q	Q I A N C C K A G V	V A A Y G Q D P A A	A V S A F Q V S V G	151
SEQ_ID_NO_127	H C C K R T P A V V	D L L P G V P Y N Q	Q I A N C C K A G V	V S A Y G Q D Q A G	S V S A F Q V S V G	154
SEQ_ID_NO_131	H C C K R T P A V V	D L L P G V P Y N Q	Q I A N C C K A G V	V S A Y G Q D P A G	S V S A F Q V S V G	157
SEQ_ID_NO_121	H C C K K T P T V V	D L L P G V P Y N Q	Q F S N C C K G G V	M A A W G Q D P T A	S V S A F Q V S V G	168
SEQ_ID_NO_123	H C C K K T P T V V	D L L P G V P Y N Q	Q F S N C C K G G V	V S S W G Q D P A A	S V S A F Q V S V G	176
SEQ_ID_NO_120	H C C K K D P T V V	D L L P G V P Y N Q	Q F S N C C K G G V	V A A W G Q D P S Q	A I S S F Q V S V G	147
SEQ_ID_NO_122	H C C K K T P T V V	D L L P G V P Y N Q	Q F S N C C K G G V	V A A W G Q D P S S	A V S S F Q V S I G	179
SEQ_ID_NO_117	H C C K K T P T V V	D L L P G V P Y N Q	Q F S N C C K G G V	I G A W G Q D P S A	A V S Q F Q V S A G	148
SEQ_ID_NO_118	H C C K K T P T L V	D L L P G V P Y N Q	Q F S N C C K G G V	V A A W G Q D P S A	S V S Q F Q V S V G	153

SEQ_ID_NO_132	N S G T T N K T V K	L P K N F T L K A P	G P G Y T C G P A K	I V K S S L F E S T	D R R R T T Q A L M	224
SEQ_ID_NO_129	L A G T T N K T V K	L P K N F T L M G P	G L G Y T C G P A T	V V P S T V Y W S A	D H R R K T Q A L M	200
SEQ_ID_NO_125	L A G T T N K T V K	L P R N F T L Q G P	G P G Y T C G P A R	I I P S T V Y L T P	D R R R R T Q A L M	201
SEQ_ID_NO_127	L A G T T N K T V K	L P K N F T L M G P	G P G Y T C G P A T	V V P S T V Y W T P	D H R R R T Q A L M	204
SEQ_ID_NO_131	L A G T T N K T V K	L P R N F T L M G P	G L G Y T C G P A A	V V P S T V Y W T P	D H R R R T Q A L M	207
SEQ_ID_NO_121	L A G T S N K T V K	L P K N F T L L G P	G P G Y T C G P A K	V V P S T V F L T P	D R R R K T Q A L M	218
SEQ_ID_NO_123	L A G T S N K T V K	L P K N F T L L G P	G P G Y T C G I A K	I V P P T N F L T P	D R R R K T Q A L M	226
SEQ_ID_NO_120	Q A G T S N K T V K	L P K N F T L F A P	G P G Y T C G P A K	I V P S T N F L T P	D K R R K T Q A L M	197
SEQ_ID_NO_122	L A G T S N K T V K	L P K N F T L L G P	G P G Y T C G P A K	V V P S T V F L T A	D K R R K T Q A L M	229
SEQ_ID_NO_117	L A G T T N K T V K	L P K N F T L L G P	G P G Y T C G P A K	I V P S T V F L T T	D K R R K T Q A L M	198
SEQ_ID_NO_118	L A G T S N K T V K	L P K N F T L L G P	G L G Y T C G P A K	I V P S T I F L T P	D K R R K T Q A L M	203

SEQ_ID_NO_132	T W N V T C T Y S Q	F L A Q K S S T C C	V S L S S F Y N D T	I T P C P T C A C A	C - - - - -	265
SEQ_ID_NO_129	T W T V T C T Y S Q	Q L A S R Y P T C C	V S F S S F Y N S T	I V P C A R C A C G	C G A H K S T G G R	250
SEQ_ID_NO_125	T W A L T C T Y S Q	Q L A S K Y P S C C	V S F S S F Y N S T	I V P C A R C A C G	C - - - - - G H	244
SEQ_ID_NO_127	T W T V T C T Y S Q	Q L A S K Y P S C C	V S F S S F Y N D T	I V P C A K C A C G	C - - - - - G H	247
SEQ_ID_NO_131	T W T V T C T Y S Q	Q L A S R Y P S C C	V S F S S F Y N S T	I V P C A R C A C G	C - - - - - G G H	251
SEQ_ID_NO_121	T W N V T C T Y S Q	F L A S K N P T C C	V S F S S F Y N E T	I T P C P T C A C G	C - - - - -	259
SEQ_ID_NO_123	T W N V T C T Y S Q	L R A R K N P S C C	V S F S S F Y N E T	I T P C P S C A C G	C - - - - -	267
SEQ_ID_NO_120	T W N V T C T Y S Q	F L A R K N P S C C	V S L S S F Y N E T	I T P C P S C A C G	C - - - - -	238
SEQ_ID_NO_122	T W N V T C T Y S Q	F L A R K N P G C C	V S L S S F Y N E T	I T P C P S C A C G	C - - - - -	270
SEQ_ID_NO_117	T W N V T C T Y S Q	F L A R K H P S C C	V S F S S F Y N D T	I T P C P S C A C G	C - - - - -	239
SEQ_ID_NO_118	T W N V T C T Y S Q	F L A R K H P S C C	V S F S S F Y N D T	I T P C P S C A C G	C - - - - -	244

ФИГ. 3 (2 из 4)

SEQ_ID_NO_132	RNNVTQPSCV	HSDSPVLKLP	GPTNPLTNSL	QPPLLQCTRH	MCPVRVHWHV	315
SEQ_ID_NO_129	GGKSHSDGCI	AGDSKRALTIP	GVNTPKKDG-	-AQLLQCTNH	MCPI RVHWHV	298
SEQ_ID_NO_125	GGH- GAAGCI	AGDSKRALSP	GVNTPRKDG-	- QPLLQCTPH	MCPVRVHWHV	291
SEQ_ID_NO_127	GGHAGPGGCI	EGDSKRALSP	GVNTPRKDG-	- QALLQCTPH	MCPVRVHWHV	295
SEQ_ID_NO_131	GGHAGPGGCI	EGDSKRALSA	GVNTPRKDG-	- QALLQCTPH	MCPI RVHWHV	299
SEQ_ID_NO_121	- - - QNKNSCV	KSNSKESHKK	GI NTPKKDN-	- TPLLQCTHH	MCPI RVHWHV	304
SEQ_ID_NO_123	- - - QNKNNCV	QSDSKLLSMV	GI NTPRKDN-	- VPLLQCTHH	MCPVRVHWHL	312
SEQ_ID_NO_120	- - - QNKKHCV	KGNSKI LSMV	GVHTPKKDN-	- EPLLQCTHH	MCPI RVHWHV	283
SEQ_ID_NO_122	- - - QNKRNCI	KSDSKRI NMV	GI HTPKKDN-	- EPLLQCTHH	MCPI RVHWHV	315
SEQ_ID_NO_117	- - - ENKKSCV	KADSKI LTKK	GLNTPKKDN-	- TPLLQCTHH	MCPVRVHWHV	284
SEQ_ID_NO_118	- - - ESKKGCV	KSDSKI LSVK	GI NTPRKDN-	- APLLQCTHH	MCPVRVHWHV	289

SEQ_ID_NO_132	KLNYKDYWRV	KI T VTNFNRY	MNYT D WTLVA	QHPN E DNVTD	VFSFNYKPLT	365
SEQ_ID_NO_129	KLNYKDYWRA	KI AVTNFNRY	MNYT QWTLVA	QHPNLNNVTE	VFSFQYKPLL	348
SEQ_ID_NO_125	KLNYKDYWRA	KI AI TNFNRY	I NYTQWTLVA	QHPNLDNVTE	VFSFQYKPLL	341
SEQ_ID_NO_127	KLNYKDYWRA	KI AI TNFNRY	MNYT QWTLVA	QHPNLDNVTE	VFSFQYKPLL	345
SEQ_ID_NO_131	KLNYKDYWRA	KI AI TNYNRY	MNYT QWTLVA	QHPNLDNVTE	VFSFQYKPLQ	349
SEQ_ID_NO_121	KVNYKDYWRA	KVAVTNFNRY	MNYTEWTLVV	QHPNLNNVTQ	VFSFDYKPLV	354
SEQ_ID_NO_123	KI NYKDYWRV	KVSI TNFNRY	LN YTLWTLVV	QHPNLNNVTQ	VFSFDYKPLV	362
SEQ_ID_NO_120	KTNYKDYWRV	KVAI TNFNRY	MNHSL WSLAV	QHPNLNNLTQ	VFSFNYKPLL	333
SEQ_ID_NO_122	KLNY M DYWRV	KVAVTNFNRY	MNYSL WTLAV	QHPNLNNVTQ	VFSFDYKPI L	365
SEQ_ID_NO_117	KTNYKDYWRV	KI AI TNFNRY	MNHTL WTLAI	QHPNLNNVTQ	VFSFDYKPV S	334
SEQ_ID_NO_118	KVNYKDYWRV	KI AVTNFNRY	MNFSL WTLAI	QHPNLNNVTQ	VFSFDYKPLV	339

SEQ_ID_NO_132	PYGSI NDTAM	FWGQKYNDL	L MQAGP MGSV	QSELLL RKDK	QTFTFKQGWA	415
SEQ_ID_NO_129	PYGNI NDTGM	FYGLK L YNDL	L MEAGPFGNV	QSEVL MRKD D	ATFTF GQGWA	398
SEQ_ID_NO_125	PYGAI NDTGM	FYGLKFYNDL	L MEAGPFGNV	QSEVL MRKDA	RTFTFSQGWA	391
SEQ_ID_NO_127	PYGSI NDTGM	FYGLKFYNDL	L MEAGPFGNV	QSEVL MRKDA	RTFTFSQGWA	395
SEQ_ID_NO_131	PYGSI NDTGM	FYGLKFYNDL	L MEAGPFGNV	QSEVL MRKDA	RTFTFS M GWA	399
SEQ_ID_NO_121	PYESI NDTGM	FYGMKFYNDL	L MEAGPFGNV	QSEVLL QKDK	NTFSL KQGWA	404
SEQ_ID_NO_123	PYESI NDTGM	FYGMKFYNDL	L MEAGPFGNV	QSEVL F QKD	NSFTFKQGWA	411
SEQ_ID_NO_120	PYESI NDTGM	FYGMKYFNDL	L MEAGPTGNV	QSELLL QKDK	D AFTFKQGWA	383
SEQ_ID_NO_122	PYESI NDTGM	FYGMKYFNDL	L MEAGPTGNV	QSEI LL QKDK	DTFTL KQGWA	415
SEQ_ID_NO_117	PYGSI NDTGM	FYGT KFYNDL	L MEAGP S GNV	QSEVLL QKDK	K TFTFKQGWA	384
SEQ_ID_NO_118	PYESI NDTGM	FYGMKYNDL	L MEAGPFGNV	QSEVLL QKDR	NTFTFKQGWA	389

ФИГ. 3 (3 из 4)

SEQ_ID_NO_132	FPRRLYFNGD	QCVMPSPDAY	PWLPSTAHPA	PAPPSLFAFI	LPLALALAI V	465
SEQ_ID_NO_129	FPRKI YFNGD	ECKMPPPPDSY	PYLPNS - - -	- APPRSSIIT	AAASTCLVLL	443
SEQ_ID_NO_125	FPRKI YFNGD	ECKMPPPPDSY	PYLPNS - - -	- AP AAPVI	ATAASAFL LV	435
SEQ_ID_NO_127	FPRKI YFNGD	ECKMPPPPDSY	PYLPNA - - -	- APVAASQLV	VSAAASAFLL	440
SEQ_ID_NO_131	FPRKI YFNGD	ECKMPPPPDSY	PYLPNA - - -	- APVVASQLV	LSAAASAFLL	444
SEQ_ID_NO_121	FPRKVYFNGD	ECMLPPPDY	PYLPNS - - -	- AYANPTSI	- LSMAASLLL	447
SEQ_ID_NO_123	FPRKVYFNGD	ECMLPPPDY	PFLPNS - - -	- AHQNL LAF	- STEI SSVIF	454
SEQ_ID_NO_120	FPRKVYFNGD	ECMLPPPDY	PFLPNS - - -	- APASLLNF	- - PAFMLLLF	425
SEQ_ID_NO_122	FPRKVYFNGE	ECMLPPPDY	P LPNS - - -	- APVNLLNF	- - TVFI L TML	457
SEQ_ID_NO_117	FPRKVYFNGD	ECMLPPPPDSY	PFLPNS - - -	- AQGNFASF	- - SLTILL L	425
SEQ_ID_NO_118	FPRKVYFNGD	ECMLPPPDY	PFLPNS - - -	- AHGNLSSLW	T LPLALLLLL	434

SEQ_ID_NO_132	PSLFLP	471
SEQ_ID_NO_129	LLLLAA	449
SEQ_ID_NO_125	ALLLVA	441
SEQ_ID_NO_127	ALLLVA	446
SEQ_ID_NO_131	LLLLVA	450
SEQ_ID_NO_121	ILL SMG	453
SEQ_ID_NO_123	LLVT VW	460
SEQ_ID_NO_120	FL LAVW	431
SEQ_ID_NO_122	VMLALW	463
SEQ_ID_NO_117	LEI SIW	431
SEQ_ID_NO_118	LLI SIW	440

ФИГ. 3 (4 из 4)

SEQ_ID_NO_10	GLLYRGGT--	-----GGR-	-----FPFQS-	---YPTA--	--APSAQQFF	154
SEQ_ID_NO_2	DLFHRAAAVT	AVAGGGMR--	-----FPFNG-	---YPVAP--	--RQPMHPYF	177
SEQ_ID_NO_1	DLFHRAAAVT	AVAGTGVR--	-----FPFRG-	---YPVA--	--RPATHPYF	166
SEQ_ID_NO_4	DLFHRAAAVT	AVSPGGMR--	-----FPFKG-	---YPVA--	--RPTPNPYF	173
SEQ_ID_NO_6	DLFHRAAAVN	AVTAGGMR--	-----FPFKG-	---YPVA--	--RPTPHQYF	172
SEQ_ID_NO_8	DLFHRAAAVN	AVSTGMR--	-----FPFKG-	---YPVA--	--CPTPQQYF	171
SEQ_ID_NO_14	TLNIPSHQ--	---HHHLR--	---HGHEPT	GVI FPGGAW	SLAAEAHPVF	175
SEQ_ID_NO_15	DLTRRLGSLVA	ADGGDNCRRS	GEVGYPI FHQ	QPTVAM--	--LPNGQPVL	160
SEQ_ID_NO_13	NLSHPAAA--	-GQFSAVR--	---YPAVG-	--VFPIA--	-----QPLF	151
SEQ_ID_NO_11	NLAPAAAA--	-----SAR--	---FPFQHQ	FPVFTGA--	--VPAANQVL	149
SEQ_ID_NO_12	NLA- PAGA--	-GFAGSI R-	---FPFQQP	FAVFPGG--	--MPAAKQAL	169

SEQ_ID_NO_10	FLD- - - - AK	SATNNR- - - -	-----	-QLALC- PPM	FVACF QPPPA	183
SEQ_ID_NO_2	FYEQAAAAAA	AASGYR- - - -	-----	-ALKVA- QPV	TVAAV- - - -	206
SEQ_ID_NO_1	FYEQAAAAAA	AEAGYR- - - -	-----	-MMKLA- PPV	TVAAV- - - -	195
SEQ_ID_NO_4	FYEQAAAAAA	AAAGYR- - - -	-----	-MLKVAPPPV	TVAAV- - - -	203
SEQ_ID_NO_6	FYEQAAAAAA	AAAGYR- - - -	-----	-MLKVAPPPV	TVAAV- - - -	202
SEQ_ID_NO_8	FYEQAAAAAA	AASGYR- - - -	-----	-MLKVAPPV	TVAAV- - - -	201
SEQ_ID_NO_14	FFDAFSVQGE	SNNNNKNNI I	NNNKI HSKNI	NLCRLDRTVM	VSSG- - - -	219
SEQ_ID_NO_15	LFDSLWRAGV	VN- - - R- - -	-----	-PQPYHVTM	GFNGV- - - NA	189
SEQ_ID_NO_13	FFEPFSRPEK	PKTHRD- - - -	-----	-MFDLDRVA	DFHPA- - - IA	183
SEQ_ID_NO_11	YFDAVL RAGM	AGPRGF- - - -	-----	-AFGYNHHPV	AASEF- - - HA	181
SEQ_ID_NO_12	YLDAVL RASM	ASHGQF- - - -	-----	-GFGYNRPAA	AAAG- - - -	198

SEQ_ID_NO_10	VAVQSDSDSS	SVV- - DLHP	DYRSPPPQAK	AFPLAFDLNL	--PPPEEIV	227
SEQ_ID_NO_2	--ARSDSDSS	SVV- - DLSP	---SPPAVTA	HKAVAFDLDL	NRPPPSSE-	246
SEQ_ID_NO_1	--AQSDSDSS	SVV- - DLAP	---SPPAVTA	NKAAAFDLDL	NRPPPVEN-	235
SEQ_ID_NO_4	--AQSDSDSS	SVV- - DRTP	---SPPAVAA	KKEVSFDLDL	NWPPPAEN-	243
SEQ_ID_NO_6	--AQSDSDSS	SVV- - DHTP	---SPPAVTA	KKEVGFDLDL	NWPPPAEN-	242
SEQ_ID_NO_8	--AQSDSDSS	SVV- - DHSP	---SPPAVTE	NKVGFDLDL	NWPPPAEN-	240
SEQ_ID_NO_14	--VHSDSDSS	SVVVDYDHDR	---SP- - - C	NKGLSLDLDL	NFPPAEVA-	258
SEQ_ID_NO_15	GVGPTVSDSS	SAV- - EENQ	-----Y	DGKRGIDLDL	NLAPPMEF-	225
SEQ_ID_NO_13	GSVHSDSDSS	SVV- - DFNY	---HD- - - -	RSTRLLNLDL	NHPPAEVA-	220
SEQ_ID_NO_11	--TTSDSDSS	SVI- - DLNH	---NEGEMKG	NGSRIFDLDL	NHPPPHI A	222
SEQ_ID_NO_12	--AQSDSDSS	SVI- - DLNQ	---NEGDMVK	NNGRGLMLDL	NEPPPQEMA	239

ФИГ. 4 (2 из 2)

SEQ_ID_NO_660	- MSLPSARLC	TVPATRNLS	SVSFSTATSL	- - - - LPKVS	V-TRGASVSL	43
SEQ_ID_NO_653	MAASP- - - -	- - AAGAAAAT	VSSFVSPSSF	SSVKASKPDR	LRPARRAAAV	43
SEQ_ID_NO_655	MAAAPASSAA	AMAAAPAI TA	SSSFVSSSPC	- SLRMSKTSP	RRPTGRVSHK	49
SEQ_ID_NO_657	MAAAP- - - -	- - AGAPAI TA	SSSFLASPPF	- ALKASTTSH	RRPAGRVSVN	42
SEQ_ID_NO_659	MAAAP- - - -	- - AGAPAI TA	SSSFVSSSPF	- CLKASMTSQ	RRPAGRVSIV	42
SEQ_ID_NO_652	MYSLM- - - -	- - SSPSI SSS	SSSFLGQTHF	- DSR- NPNAC	L-PT- KDSM	39
SEQ_ID_NO_651	MSSTH- - - -	QLVSSMI SSS	SSTFLAPSNF	- NL R- TRNAC	L- PMAKRVNT	42
SEQ_ID_NO_649	MSPTQ- - - -	LLSTSISSSS	SAFLAAPIAF	- KAR- NQNV	I- AS- KTPSI	41
SEQ_ID_NO_645	MSITH- - - -	- SLTTPLS	SSAFLAPSSF	- NCRGQVSLP	V- - - KSVSI	39
SEQ_ID_NO_647	MSLMN- - - -	NLSTSMI SSS	SSTFLAPTST	- NSSRTQSVS	V- PL- KSI NI	42
SEQ_ID_NO_660	- I RCI AEP LA	EKTTYTTSVN	RNANI AKLQA	GYLFPEI ARR	RNAHI QRYPD	91
SEQ_ID_NO_653	NVRCVSSPPA	TETSFKTKVP	RNANMAKLQA	GYLFPEI ARR	RAAHLLKFPD	93
SEQ_ID_NO_655	- I SCVSSPPA	AETSYKTSVP	RNANMAKLQA	GYLFPEI ARR	RAAHLLKFPD	98
SEQ_ID_NO_657	- I RCVSSPPA	VDTSYKTNVP	RNANMAKLQA	GYLFPEI ARR	RAAHLLKYPD	91
SEQ_ID_NO_659	- I RCVSSPPA	VDTSFKTNVP	RNANMAKLQA	GYLFPEI ARR	RAAHLLKFPD	91
SEQ_ID_NO_652	- CKCVATPST	EITTAHTTKVS	RNANMAKLQA	GYLFPEI ARR	SAHMLKYPD	88
SEQ_ID_NO_651	- CKCVATP Q	EKL EYKTKVS	RNSNMSKLQA	GYLFPEI ARR	SAHLLKYPD	90
SEQ_ID_NO_649	- CTC A A A P Q E	QKT V YKTQVS	RNANI AKLQA	GYLFPEVARR	RNAHMLKYPD	90
SEQ_ID_NO_645	- CKCVATP E	AETAYKT GVN	RNPNM GKLQA	GYLFPEI ARR	SAHLLKYPD	87
SEQ_ID_NO_647	- VKCVATPQE	QQNAYKTKVS	RNANI AKLQA	GYLFPEVARR	RAAHLLKYPN	91
SEQ_ID_NO_660	AKVI SLGI GD	TTEPIPTVIT	GAMEARARAL	STLEGYSGYG	AEQGEKPLRA	141
SEQ_ID_NO_653	AKI I SLGI GD	TTEPI PDVIT	NAMAKRAHAL	STVDGYSGYG	AEQGEKKLRA	143
SEQ_ID_NO_655	AKI I SLGI GD	TTEPI PDVIT	NAMAERAHAL	STI DGYSYG	AEQGEKKLRS	148
SEQ_ID_NO_657	AKI I SLGI GD	TTEPI PNVIT	NAMAERALAL	STI DGYSYG	AEQGEKKLRA	141
SEQ_ID_NO_659	AKI I SLGI GD	TTEPI PNVIT	NAMAERAHAL	STI DGYSYG	AEQGEKKLRA	141
SEQ_ID_NO_652	AQVI SLGI GD	TTEPI PEVIT	SGMAKKAHAL	STLEGYSGYG	AEQGEKQLRA	138
SEQ_ID_NO_651	AQVI SLGI GD	TTEPI PEVIT	SAMAKKAHEL	STI EGYSGYG	AEQGA KPLRA	140
SEQ_ID_NO_649	AKVI SLGI GD	TTEPI PEVIT	SAI AKRAEAL	STLEGYSGYG	PEQGEKPLRT	140
SEQ_ID_NO_645	AKVI SLGI GD	TTEPI PEVIT	DAMSKRSHAL	STI EGYSGYG	AEQGEKPLRR	137
SEQ_ID_NO_647	AQVI SLGI GD	TTEPI PDVIT	SAMAKRSHAL	STLEGYSGYG	AEQGEKALRA	141

ФИГ. 5 (1 из 4)

SEQ_ID_NO_660	GI GAIFYADL	GI DETEI FVS	DGAKCDI TRL	QLVFGPNVTM	AAQDPSYPAY	191
SEQ_ID_NO_653	AI AATYYADL	GI EETDI FVS	DGAKCDI SRL	QVLFGSNVKI	AVQDPSYPAY	193
SEQ_ID_NO_655	AI AATYYVDL	GI ESDDI FVS	DGAKCDI SRL	QVLFGSNVTI	AVQDPSYPAY	198
SEQ_ID_NO_657	AI AATYYADL	GI ESDDI FVS	DGAKCDI SRL	QVLFGSNVTI	AVQDPSYPAY	191
SEQ_ID_NO_659	AI AATYYADL	GI ESDDI FVS	DGAKCDI SRL	QVLFGSNVTI	AVQDPSYPAY	191
SEQ_ID_NO_652	AI ASTFYGDL	SI EESDI FVS	DGAKSDI SRL	QVMFGSNVTM	AVQDPSYPAY	188
SEQ_ID_NO_651	AI AKIFYGGL	GI GDDDV FVS	DGAKCDI SRL	QVMFGSNVTI	AVQDPSYPAY	190
SEQ_ID_NO_649	AI ASTFYSGL	GI EEDDI FVS	DGAKCDI SRL	QMVFGANVTM	AVQDPSYPAY	190
SEQ_ID_NO_645	ALASTFYSDL	GI EEDDI FVS	DGAKCDI SRL	QI VFGSNVKM	AVQDPSYPAY	187
SEQ_ID_NO_647	ALASTFYFNL	GI EDDDI FVS	DGAKCDI SRL	QVVFSGSNVTM	AVQDPSYPAY	191
SEQ_ID_NO_660	VDTSVMMGQT	GLFQSDSQY	SKI QYMKCTP	ENDFFPDLSS	TPRTDI I FFC	241
SEQ_ID_NO_653	VDSSVI MGQT	GLYQEDVQKY	GNI EYMKCSP	ENGFFPDLSS	VPRTDI I FFC	243
SEQ_ID_NO_655	VDSSVI MSQT	GLYQQDVQKY	GNI EYMR CNP	ENGFFPDLST	VPRTDI I FFC	248
SEQ_ID_NO_657	VDSSVI MGQT	DL YQQDVQKY	GNI QYMR CSP	ENGFFPDLST	I PRTDI I FFC	241
SEQ_ID_NO_659	VDSSVI MGQT	DL YQQDVQKY	GNI EYMR CQP	ENGFFPDLST	VPRTDI I FFC	241
SEQ_ID_NO_652	VDSSVI LGQT	GQFQKDVEKY	GNI EYMK CNP	ENGFFPDLST	VSRTDI I FFC	238
SEQ_ID_NO_651	VDSSVI MGQT	GQFNTDVQKY	GNI EYMRCTP	ENGFFPDLST	VGRTDI I FFC	240
SEQ_ID_NO_649	VDSSVI MGQT	GQFQKDVEKY	GKI EYMRCTP	ENGFFPDL SK	VSRTDI I FFC	240
SEQ_ID_NO_645	VDSSVI MGQT	GLFQKNVEKF	ANI EYMR CNP	ENGFFPDLSS	ISRPDI I FFC	237
SEQ_ID_NO_647	VDSSVI MGQT	GQFQKDVEKY	GNI EYMRCTP	ENGFFPDL SK	VARTDI I FFC	241
SEQ_ID_NO_660	SPNNPTGASA	SRKQLEELVA	FAKKNCSI I V	YDSAYAI YTS	DDSPKSI YEI	291
SEQ_ID_NO_653	SPNNPTGAAA	SRDQLTKLVK	FAKDNGSI I V	YDSAYAMYI S	DDSPKSI FEI	293
SEQ_ID_NO_655	SPNNPTGAAA	SRDQLTRLVK	FAKDNGSI I V	YDSAYAMYI S	DDSPKSI FEI	298
SEQ_ID_NO_657	SPNNPTGAAA	SRDQLTKLVK	FAKDNGSI I V	YDSAYAMYI S	DDSPKSI FEI	291
SEQ_ID_NO_659	SPNNPTGAAA	SRDQLTKLVK	FAKDNGSI I V	YDSAYAMYI S	DDSPKSI FEI	291
SEQ_ID_NO_652	SPY NPTGNAA	TREQLTRLVQ	FAKDNGSI LV	YDSGYAMYI S	DDSPRSI FEI	288
SEQ_ID_NO_651	SPNNPTGAAA	TREQLTQLVE	FAKKNCSI I V	YDSAYAMYMS	DDNPRS I FEI	290
SEQ_ID_NO_649	SPNNPTGSAA	TREQLTQLVQ	FAKDNGSI I V	YDSAYAMYMS	DDNPRS I FEI	290
SEQ_ID_NO_645	SPNNPTGAVA	TREQLTQLVQ	FAKDNGSI VI	HDSAYAMYI S	GDNPRS I FEI	287
SEQ_ID_NO_647	SPNNPTGAAA	TREQLTRLVK	FAKDNGSI I V	YDSAYAMYMS	DDNPRS I FEI	291

ФИГ. 5 (2 из 4)

SEQ_ID_NO_660	PGAKECAI ET	ASFSKYAGFT	GVRLGWT VVP	KALKFADGHP	VHTDFNRVMT	341
SEQ_ID_NO_653	PGAKEVAI ET	ASFSKYAGFT	GVRLGWT VVP	KELLFSDGHP	VAKDFNRI VC	343
SEQ_ID_NO_655	PGAREVAI ET	ASFSKYAGFT	GVRLGWT VVP	KELLFSDGHP	VAKDFNRI VC	348
SEQ_ID_NO_657	PGAKEVALET	ASFSKYAGFT	GVRLGWT VVP	KELLFSDGHP	VAKDFNRI VC	341
SEQ_ID_NO_659	PGAKEVAI ET	ASFSKYAGFT	GVRLGWT VVP	KELLFSDGHP	VAKDFNRI VC	341
SEQ_ID_NO_652	PGAKEVAI EV	SSFSKYAGFT	GVRLGWT VVP	KELLYSDGFP	VAKDFNRI EC	338
SEQ_ID_NO_651	PGAEEVAMET	ASFSKYAGFT	GVRLGWT VI P	KKLLYSDGFP	VAKDFNRI IC	340
SEQ_ID_NO_649	PGAKEVALET	SSFSKYAGFT	GVRLGWT VVP	KQLLYSDGFP	VAKDFNRI VC	340
SEQ_ID_NO_645	PGAKEVAI ET	SSFSKYAGFT	GVRLGWT VVP	KQLLFSDGFP	VAKDFNRI VC	337
SEQ_ID_NO_647	PGAKEVAI ET	ASFSKYAGFT	GVRLGWT VI P	KQLLFSDGFP	VAKDFNRI VC	341

SEQ_ID_NO_660	TCFNGASNVA	QAGGLACVSS	EGLKAMHETV	KFYKENTKI L	VETFE SLGFK	391
SEQ_ID_NO_653	TCFNGASNI S	QAGGLGCL SP	EGLKAMSDVV	GFYKENTKI I	VDTFTSLGFN	393
SEQ_ID_NO_655	TCFNGASNI A	QAGGLACL SP	EGLKAMHDVV	GFYKENTEI I	VDTFTSLGFN	398
SEQ_ID_NO_657	TCFNGASNI S	QAGGLACL SP	EGLKAMHDVV	GFYKENTEI I	VDTFTSLGFN	391
SEQ_ID_NO_659	TCFNGASNI A	QAGGLACL SP	DGLKAMQDVV	GFYKENTEI X	VETXTSLGFN	391
SEQ_ID_NO_652	TIFNAASNI S	QAGGLACL SP	EGL EAMHKL V	GFYKENTNI I	METFTSLGFS	388
SEQ_ID_NO_651	TCFNGASNI S	QAGGLACL TP	EGL EAMHKVI	GFYKENTNI I	IDTFTSLGYD	390
SEQ_ID_NO_649	TIFNGASNI C	QAGGLACL SP	EGLKAMSEVI	GFYKENSNI I	MDTFNSLGFN	390
SEQ_ID_NO_645	TCFNGASNI S	QAGGLACL SP	EGLKAMRDVI	GFYKENTNI I	METFD SLGFK	387
SEQ_ID_NO_647	TCFNGASNI A	QAGGLACL SS	EGL EAMQEV I	GFYKENTKI I	VETFNSLGFK	391

SEQ_ID_NO_660	TFGGKNAPYV	WWQFPGKSSW	DVFSEI LEQT	HI VTTPGSGF	GPGGEGFI RA	441
SEQ_ID_NO_653	VYGAKNAPYV	WHFPGRNSW	DVFAEI LEKA	HVVTTPGSGF	GPGGEGFVRV	443
SEQ_ID_NO_655	VYGAKNAPYV	WHFPGRNSW	DVFAEI LEKA	NVTTTPGSGF	GPGGEGFVRV	448
SEQ_ID_NO_657	VYGAKNAPYV	WHFPGRNSW	DVFAEI LEKA	NVTTTPGSGF	GPGGEGFVRV	441
SEQ_ID_NO_659	VYGAKNAPYV	WHFPGRNSW	DVFAEI LEKA	NVTTTPGTGF	GPGGEGFVRV	441
SEQ_ID_NO_652	VYGGKNAPYV	WHFPQGSSW	DVFSEI LEKT	HVVTTPGSGF	GPA GDGFI RV	438
SEQ_ID_NO_651	VYGGKNAPYV	WHFPNQSSW	DVFAEI LEKT	HVVTTPGSGF	GPGGEGFVRV	440
SEQ_ID_NO_649	VYGGKNAPYV	WHFPQGSSW	DVFSEI LEKT	HVVTTPGSGF	GPGGEGFVRV	440
SEQ_ID_NO_645	VYGGKDAPYV	WHFPGRSSW	DVFAEI LEKT	HVVTTPGSGF	GPGGEGFI RV	437
SEQ_ID_NO_647	VYGGKNAPYV	WHFPGRSSW	DVFSEI LEKT	HI VTTPGSGF	GPGGEGFI RV	441

ФИГ. 5 (3 из 4)

SEQ_ID_NO_660	SAFGHRENI L	EASRRRLKEMF	GSKK	465
SEQ_ID_NO_653	SAFGHRENI I	EAARRLKQLY	K - -	464
SEQ_ID_NO_655	SAFGHRENI I	EAARRLKQLY	K - -	469
SEQ_ID_NO_657	SAFGHRDNI I	EAARRLKQLY	K - -	462
SEQ_ID_NO_659	SAFGHRENI I	EAARRXKQLY	K - -	462
SEQ_ID_NO_652	CAFSTRGNVL	EACKRFKRLY	K - -	459
SEQ_ID_NO_651	SAFGHRENI L	EACRRFKQLY	K - -	461
SEQ_ID_NO_649	SAFGHRENVL	EACRRFKQLY	N - -	461
SEQ_ID_NO_645	SAFGHRENVL	EACRRFKQLY	K - -	458
SEQ_ID_NO_647	SAFGHRENVL	EACRRFKQLY	N - -	462

ФИГ. 5 (4 из 4)

SEQ_ID_NO_257

SEQ_ID_NO_253

SEQ_ID_NO_255

MAF

GVR

LCL

CLLL

VFA

VT

S

ARNT

SFS

DNE

MA

LA

KG

RSL

K

TL

NDY

MTF

VVR

LLV

CLLL

TLT

ITS

SL

ARNP

VSVS

GGF

ENS

GF

QR

SL

L

MVN

VEDY

MSFV

VLRL

AV

FLLL

TLT

VTY

S

-

-

SP

SVS

VPV

V

K

I

G

IER

RSLI

VNV

KDY

48

49

47

SEQ_ID_NO_257

SEQ_ID_NO_253

SEQ_ID_NO_255

GDPI

ANR

GHD

PS

QRN

KN

WG

SGG

GR

K

G

GDPS

ANP

KHD

PGV

PPS

AT

GQ

RVV

GR

G

DGP

SANP

KHN

PGT

PPV

TS

Q

RSP

GR

G

75

75

73

ФИГ. 6

SEQ_ID_NO_334	MASLSMTSS	SVRAQWTKKQ	NKLFEQALAV	YDKDTPDRWH	NIARAVGGK	48
SEQ_ID_NO_336	MASMSLS--	SSRAQWTAKQ	NKLFEQALAV	YDRDTPDRWH	NIARAVGGK	46
SEQ_ID_NO_327	MASSSMS--S	QSSGSWTAKQ	NKAFEQALAT	YDQDTPNRWQ	NVAKVGGK	47
SEQ_ID_NO_332	MSSMSSQ-H	GSSGSWTAKQ	NKAFEKALAV	YDKETFDWSS	NVAKAVGGK	47
SEQ_ID_NO_329	MASSSI--	-ASGSWSVKD	NKAFEKALAV	YDKDTPDRWY	NVAHAVGGK	45
SEQ_ID_NO_337	MASSSMS--	-ASGSWSVKE	NKAFERALAV	YDKDTPDRWY	NVAHAVGGK	45
SEQ_ID_NO_326	-MSSSYQASR	NSRSAWTPRE	NKLFEKALAL	FDKDTPDRWQ	NIKAVGGVK	49
SEQ_ID_NO_331	-MASSITLSR	DSNSYMTPKQ	NKLFERALAV	YDKDTPDRWQ	KVAAGVE-K	48
SEQ_ID_NO_323	MASSSLKQK	ASDSSWTPKQ	NKLFEKALAK	YDKDTPDRWQ	NVAKAVGGK	49
SEQ_ID_NO_324	MASGSRM--	-GSGLWTSKQ	NKLFEKALAL	YDKDTPDRWQ	NVAKAVGGK	45

SEQ_ID_NO_334	SAEEVRRYYE	MLEEDVKHIE	SGKVPLPAYR	CPGGAGAGAL	GYEADRMKHL	98
SEQ_ID_NO_336	SADEVRRYYE	LLVKDLEHIE	AGKVAFPAYR	CPGGYD---	DADSDRLKHL	92
SEQ_ID_NO_327	TTEEVKRHYE	LLVQDINSIE	NGHVPFPNYR	TSGGCTNGRL	SQEEKRMRNM	97
SEQ_ID_NO_332	TAEVVKRHYE	ILLRDVFFID	NGMVPFPKYK	TTGGSHNSTS	D-----	88
SEQ_ID_NO_329	TPEEVKRHYE	LLVQDVKHIE	SGRVFPFNYK	KTTSGST---	DQEEKRLRNL	92
SEQ_ID_NO_337	TPEEVKKHYE	LLVEDIKHIE	SGKVFPFNYK	KLSVSH---	--EEKRMRNM	89
SEQ_ID_NO_326	SAEEVKKHYE	ILLIEDLQHIE	SGRIPIPKYK	SSGSCNNT--	NEEESEVNKN	97
SEQ_ID_NO_331	SVEEVRRHYE	ILLVRDLMYIE	SGQIPIPNYK	STGSNRR---	-----	85
SEQ_ID_NO_323	SADEVKRHYE	ILLIEDLRHIE	SGHVPLPKYK	STGSSTNV--	EEEEERLLKYL	97
SEQ_ID_NO_324	SAEEVKRHYE	ILLIEDLKHIE	SGHVPIPNYK	STGSNSIG-	DQEQRLLKCI	93

SEQ_ID_NO_334	KI-	-----	-----	100
SEQ_ID_NO_336	T-	-----	-----	93
SEQ_ID_NO_327	RLQ-	-----	-----	100
SEQ_ID_NO_332	-	-----	-----	88
SEQ_ID_NO_329	NLNLQ-	-----	-----	97
SEQ_ID_NO_337	SLH-	-----	-----	92
SEQ_ID_NO_326	KLQMATGNLH	IDSLRVLQDF	MAGCNQH	124
SEQ_ID_NO_331	-	-----	-----	85
SEQ_ID_NO_323	KLN-	-----	-----	100
SEQ_ID_NO_324	KLQ-	-----	-----	96

ФИГ. 7

SEQ_ID_NO_630	-----	-MQSKHYMVV	HGMSHGAWCW	YKLKPLLESA	GHRVTALDMG	39
SEQ_ID_NO_595	-----	MSE EKRKQHFVLV	HGACHGAWCW	YKVKPLLEAL	GHRVTALDLA	43
SEQ_ID_NO_615	MNYGEGRGDG	SONKKH I LV	HGACHGAWSW	HKVTTQLRSA	GYQVTV PDLA	50
SEQ_ID_NO_623	-----	MHSAAN AKQQKHFVLV	HGGC GAW I V	YKLKPLLESA	GHKVTAVDLS	46
SEQ_ID_NO_642	-----	MEVMKHFVTV	HGVGHGAWWY	YKLKPRI EAA	GHRCTAVNLA	40
SEQ_ID_NO_611	-----	MSPTKHFVAV	HGVGHGAWWY	YKLKPRI EAA	GFKETAIDL A	40
SEQ_ID_NO_605	-----	MEEVV GMEEKHFVLV	HGVN HGAWCW	YKLKARLVAG	GHRVTAVDLA	45
SEQ_ID_NO_620	-----	MKEGKHFVLV	HGACHGGWSW	YKLKPLLEAA	GHKVTALDLA	40
SEQ_ID_NO_632	-----	MEAN KKQGKHFVLV	HGAGHGAWCW	YKLKPLLEAA	GHKVTALDLA	44
SEQ_ID_NO_622	-----	MVE TKNQKHFVLV	HGACHGAWCW	QKFKTLLESA	SNRVTVLDLA	43
SEQ_ID_NO_616	-----	MEVD RKQGRHFVLV	HGACHGAWSW	YKVKPRLEAA	GHRVTALDMA	44

SEQ_ID_NO_630	ASGVN-	MRPV	EELRSFRDYN	APLLSFMSSL	PEDDKVVLVG	HSLGGI NI AF	88
SEQ_ID_NO_595	ASGI DT	TISI	TDI STCEQYS	EPLMQLMTSL	PNDEKVVLVG	HSFGGLSLAL	93
SEQ_ID_NO_615	ASGVD-	ERRF	QDLRSFI HYS	QPLLDI LAOL	PPGERVI LVG	HSLGGLNI AL	99
SEQ_ID_NO_623	AAGI N-	PRRL	DEI HTFRDYS	EPLMEVMASI	PPDEKVVLVG	HSFGGMSLGL	95
SEQ_ID_NO_642	ASGI N-	EKKL	EEVRSS I DYA	APLLEVLDSV	PENEKVI LVG	HSGGGMTAAV	89
SEQ_ID_NO_611	AAGVN-	PKKL	EEVNSLEEYC	GPLEDVLA AV	PEGEKVI LVG	HSGGGLSA AV	89
SEQ_ID_NO_605	ASGI N-	MKRI	EDVHTFHAYS	EPLMEVLASL	PAEEKVI LVG	HSLGGVTLAL	94
SEQ_ID_NO_620	ASGT D-	LRKI	EELRTL YDYT	LPLMELMESL	SADEKVI LVG	HSLGGMNLGL	89
SEQ_ID_NO_632	ASGI D-	LRKI	EQLHTLHDYT	LPLLELMESL	PQEEKAI LVG	HSLGGMNLAL	93
SEQ_ID_NO_622	ASGAN-	MKA I	QDVE TLDEYT	EPLLEFLASL	QPK EKVI LVG	HSLGGLSLAL	92
SEQ_ID_NO_616	ASGI N-	RKQI	QEVHSMHEYS	QPLLEMMATL	PPNEKVI LVG	HSLGGLNLAV	93

SEQ_ID_NO_630	AMEEFPEKVS	AAVFVAALVP	DTVNKPSFFL	DEL FKKI -	GA	ANGWLD CQFS	137
SEQ_ID_NO_595	AMDKFPDKI S	VSVFVTA FMP	DTKHSPSFVE	EKFASS - -	MT	PEGWMGSELE	141
SEQ_ID_NO_615	AMDRFPEKI A	AAVFVTALMP	DSVNP PSYVM	DKL KKE -	KT	MLFMSDTQFG	147
SEQ_ID_NO_623	AME TYPEKI S	VAVFMSA M MP	DPNHS LTYPF	EKYNEK -	CP	ADMMLDSQFS	143
SEQ_ID_NO_642	GMEKFPPKI S	LAVFLNAI MP	DTENRPSYVL	EEYTAK -	TP	PEAWKDCQFS	137
SEQ_ID_NO_611	GMEKFPPKI S	VAVFLNAI MP	DTKNRPSYVM	EEYTAR -	TP	IEAWKDTQFS	137
SEQ_ID_NO_605	AGDKFPKI S	VAVFVTA FMP	DTTHRPSFVL	EQYSEKMGKE		DDSWLDTQFS	144
SEQ_ID_NO_620	AMEKYPQKI Y	AAVFLAA FMP	DSVHNS SFVL	EQYNER -	TP	AENWLD TQFL	137
SEQ_ID_NO_632	AMEKYPKKI Y	AAVFLAA FMP	DSI H LSSYVM	DQYNER -	TP	AENWLD TQFL	141
SEQ_ID_NO_622	AMEKFPEKI A	VAVFLSA FMP	DTTHKPSFVL	DQYNER -	TP	ADSWLDTQFL	140
SEQ_ID_NO_616	AMEKFPEKVS	VAVFLTA FMP	DTLHRPSYVL	DQYNER -	TP	NDAWLD TQFS	141

ФИГ. 8 (1 из 2)

SEQ_ID_NO_630	TFGSPDEPVT	VI SFGPKFLS	LLLYDS SPI E	DYELAKMLTR	PLPNYVTDLG	186
SEQ_ID_NO_595	TYGS DNSGL	SVFFSTDFMK	HRLYQLSPVE	DLELGLLLKR	PSSLFI NELS	190
SEQ_ID_NO_615	LVGDEDKGPV	SLLFGPKFLS	-KLYTRSPPE	DLTLARTLMR	PSSFLEDLG	196
SEQ_ID_NO_623	TYGNPENPGM	SM LGPQFMA	LKMFQNC SVE	DLELAKMLTR	PGSLFFQDLA	193
SEQ_ID_NO_642	AYG - DPPI T	SLVCGPEFIS	STLYHLSPI E	DHALGKI LVR	PGSLFIEDLL	185
SEQ_ID_NO_611	AYG - EPPI T	ALLCGPEFIS	TSLYHLSPVE	DHTLGKLLVR	PGALFVEDLL	185
SEQ_ID_NO_605	QCDASNPSHI	SMLFGREFLT	IKIYQLCPPE	DLELAKMLVR	PGSMFI DNLS	194
SEQ_ID_NO_620	PYGSPPEEPLT	SMFFGPKFLA	HKLYQLCSPE	DLALASSLVR	PSSLFMEDLS	187
SEQ_ID_NO_632	PYGTPEEPLT	SMTFGPKFLA	DKLYRLSPPE	DVALGLSLVR	TSSLFLEDLS	191
SEQ_ID_NO_622	PYSSSQSHLT	TMSFGPKFLS	SKLYQLSPPE	DLEQAKTMVR	PGSLFLYDLS	190
SEQ_ID_NO_616	PYGSSEKPN	SMFFGPEFIS	TKLYQLSPI E	DLELMLALAR	PASLFLEDLA	191

SEQ_ID_NO_630	K AEKLS DGK	YGSVRRVYVI	CKEDKAI PDE	LVGQMI EWN	GLKEVI ELQG	234
SEQ_ID_NO_595	K MENFSEKG	YGSVPRAYI V	CKEDNI I SED	HQRWMI HNY	PANLVI EME	238
SEQ_ID_NO_615	S MPPFSESG	YGSVEKI YVV	CAQDEI LTEG	FQRWMI ENN	PVKEVRELED	244
SEQ_ID_NO_623	K AKKFSTER	YGSVKRAYI E	CNEDKS FPE	FQKW E VESV	GADKVKKEI KE	241
SEQ_ID_NO_642	K AEKFTEEG	FGSVPRVYVI	AAEDKTI PPE	FQRWMI ENN	PVKEVKEI KG	233
SEQ_ID_NO_611	KGAVKFTDEG	FGSVPRVYVV	ATEDKTI PPE	FQRWMI ENN	PVAEVKEI EG	234
SEQ_ID_NO_605	K ESKFSDG	YGSVKRVYLV	CEEDIGLPKQ	FQHMMI QNY	PVNEVMEI KG	242
SEQ_ID_NO_620	K AKYFTDER	FGSVKRVYI V	CTEDKGI PEE	FQRWQI DNI	GVTEAI EI KG	235
SEQ_ID_NO_632	K AKYLTDEG	YGSVKRVYVV	CTEDKGI SKE	FQQWQI DNI	GVTEAKEI KG	239
SEQ_ID_NO_622	K ANSFSTTG	YGSVKRVYVI	CDEDLAI PEE	FQRWMI ENS	AVEEVMEI EG	238
SEQ_ID_NO_616	E LKKFSNEG	YGSVTSVFIR	CDKDEAI RKE	FQQWMI ENSG	GVKEVMNI KD	240

SEQ_ID_NO_630	ADHMPML SNP	QQLCDCLVQI	AVENP	259
SEQ_ID_NO_595	T DHMPMFCKP	QVLSDHLLAI	ADNFS	263
SEQ_ID_NO_615	ADHMPMFSTP	KQLFQCLSDV	ADACA	269
SEQ_ID_NO_623	ADHMGMLSQP	REVCKCLLDI	SDS - -	264
SEQ_ID_NO_642	ADHMPMF SKP	DELSQCCLLDI	AKKHA	258
SEQ_ID_NO_611	ADHLPQF SKP	DELTOVLVDI	AKNHG	259
SEQ_ID_NO_605	GDHMAMLSDP	QKLCDCLSQI	SLKYA	267
SEQ_ID_NO_620	ADHMAMLC EP	QKLCA SLLEI	AHKYN	260
SEQ_ID_NO_632	ADHMAMLCMP	KKLCDTLVEI	ADKYN	264
SEQ_ID_NO_622	ADHMVMF SKP	QELFHCLSEI	ANKHA	263
SEQ_ID_NO_616	ADHMAMF SKP	EELCACLLEV	AHKYG	265

ФИГ. 8 (2 из 2)

SEQ_ID_NO_88	MAASTLQYHF	AFFLFFLVIA	LI TR- - - - S	NPPPPQPKC- A	MME- - - AMGE	24
SEQ_ID_NO_86	-MAASSLNLL	LILSLLTFIS	SFSPI- - - Q	AQVDETSL- K	AMRDALEWPM	46
SEQ_ID_NO_77	MAAAPLTLL	I ALSLL- FIS	LQRS- - - - E	SLSDNPSL- T	LLPDGFDWPI	43
SEQ_ID_NO_79	-MESPI LFFL	I I FLFTITNN	LQRS- - - - E	SLPENPSM- I	LLGDGFDWPI	43
SEQ_ID_NO_87	MACSSLPKLL	VLCAFFVYIS	AI VI- - - - E	AQVDKFGLEQ	LVSEDFELPM	44
SEQ_ID_NO_82	MASSSVLSKL	LLP- - - LFLT	NVVV- - - - V	AQVDESSL- K	LI TDSLEWPS	44
SEQ_ID_NO_81	MASLLTLALL	LFC- - - IAFIS	HI AVS- SKTE	AQVEDTRL- K	LVRDAFEWPL	46
SEQ_ID_NO_84			HVAV- - - - Q	AHMEDTAF- N	LMSDALEWPT	42

SEQ_ID_NO_88	CGA- - - - -	- - - - - NV	DEENENEVGH	GR- - LLRRI R	YYI SYGALAA	57
SEQ_ID_NO_86	SMY- YDESSG	LNDGLVGFDD	GVVDDEESSR	RS- LFWRRTH	YYI SYGALSA	94
SEQ_ID_NO_77	SHS- - - DEFD	I IDGEESFEV	TEEDDGVTDR	RS- LYWRRTK	YYI SYGALSA	89
SEQ_ID_NO_79	SHG- - - DALD	I DGGES- - FE	DAEEEDVADR	RS- LYWKRRR	YYI SYGALSA	87
SEQ_ID_NO_87	AMS- - - SLYE	E- - - - - TE	DDEMQLDGNG	RS- LLWHKFK	YYI SYGALSA	83
SEQ_ID_NO_82	TMS- - - - LYN	- - EFGD- - ED	GEDPDGVIDR	RS- MFWHRMR	YYI SYGALSA	85
SEQ_ID_NO_81	SMSFQGDLHD	NEEGDE- - EV	DDEVENGYSG	RSL LWWKMR	YYI SYGALSA	94
SEQ_ID_NO_84	TMS- - - - LYD	- - - - ED- - DA	QEDVENAYSRI	RS- LFWRRMK	YYI SYGALSA	81

SEQ_ID_NO_88	NRI PCPPRSG	RSYYTHNCYR	ATEPVRPYHR	SCTAI TRCLR	DTS	100
SEQ_ID_NO_86	NRI PCPA RSG	RSYYSHNCFA	SRA PVNPYSR	GCSRI ARCR	- - -	134
SEQ_ID_NO_77	NRVPCPPRSG	RSYYTHNCFR	ARGPVHPYSL	GCSSI TPGDR	- - -	129
SEQ_ID_NO_79	NRVPCPPRSG	RSYYTHNCFR	ARGQVHPYSR	GCSSI TRCRR	- - -	127
SEQ_ID_NO_87	NRI PCPPRSG	RSYYTHHCYH	ATGPAHPYTR	GCSAI TRCRR	- - -	123
SEQ_ID_NO_82	NRI PCPPRSG	RSYYTHNCFQ	ARGPVRPYTR	GCS T TRCRR	- - -	125
SEQ_ID_NO_81	NRI PCPPRSG	RSYYTHNCFK	AHGPVHPYTR	GCSRI TRCRR	- - -	134
SEQ_ID_NO_84	NRI PCPPRSG	RSYYTHNCYR	ARGPVHPYSR	GCSAI TRCRR	- - -	121

ФИГ. 9

SEQ_ID_NO_210	-----MSA	EMVVGQGGSS	-----	-----	-----	13
SEQ_ID_NO_209	MSSDRHTPTK	DPPDHPSSSS	-----	-----	-----	20
SEQ_ID_NO_230	MSNDQIIIEG	EGGGGGEGGS	-----	-----	-----	20
SEQ_ID_NO_214	MSGDQ--RPK	DSAEGSSRSG	-----	-----	-----	18
SEQ_ID_NO_239	MSNDF--RKR	DSAEGSSRST	-----	-----	-----	17
SEQ_ID_NO_212	MSSE--RGK	DI AEGPSRVN	-----	-----	-----	17
SEQ_ID_NO_226	MSS--K GK	EI AEGSSITTT	-----	-----	-----	16
SEQ_ID_NO_220	MESTD--SGS	QQHGGDPGPS	-----	-----	-----	18
SEQ_ID_NO_236	ME-----PGP	DAPRGE GASA	-----	-----PEESGP	SSSSMAVEKA	31
SEQ_ID_NO_238	ME-----PGP	DAPAGGGGGG	GTSI SEPAEA	GPSPSSSSSAA	AAASSSSRQQ	45
SEQ_ID_NO_232	MDPSG--PGP	SSAAAGGAPA	-----	-----	-----	18
SEQ_ID_NO_242	MDPSG--PGP	SSAAAGGAPA	-----	-----	-----	18
SEQ_ID_NO_241	MDPSG--PGP	SPVVGGGGGG	-----	-----	-----	18
SEQ_ID_NO_234	MDPSG--PGP	SSVMGAAGGG	-----	-----	-----	18

SEQ_ID_NO_210	-----ATAAQLAPSS	SSNGINEGSR	ELAASFAPSR	YESQKRRDWN	53	
SEQ_ID_NO_209	-----	---NHHKQPL	PPQPQQPLSR	YESQKRRDWN	47	
SEQ_ID_NO_230	-----S SRSKTTLLIA	PSDDHHHHHQ	LPPVPPQLSR	YESQKRRDWN	61	
SEQ_ID_NO_214	-----	-----GDH	YQPQPAPLSR	YESQKRRDWN	41	
SEQ_ID_NO_239	-----	-----GE	PQQQPQPLSR	YESQKRRDWN	39	
SEQ_ID_NO_212	-----	-----PG	DQQLPPTPSR	YESQKRRDWN	39	
SEQ_ID_NO_226	-----T ATRSSHGGDD	HHHHHNQQQQ	QQQQQLPLSR	YESQKRRDWN	57	
SEQ_ID_NO_220	-----	-----SVTP	SSPPATPPSR	YESQKRRDWN	42	
SEQ_ID_NO_236	EEPPQAQAQPP	EGGGQQVALQ	EHLQPQPLSQ	QPPVPAGLSR	YESQKRRDWN	81
SEQ_ID_NO_238	AEQEAAQQQQG	AQQREQPAVR	AQAQPQPQPH	QQQPAPGLSR	YESQKRRDWN	95
SEQ_ID_NO_232	-----	-----VAA	APQPPAQLSR	YESQKRRDWN	41	
SEQ_ID_NO_242	-----	-----VAA	APQPPAQLSR	YESQKRRDWN	41	
SEQ_ID_NO_241	-----	-----DAPAVV	PQQRPAQLSR	YESQKRRDWN	44	
SEQ_ID_NO_234	-----	-----EAPAV	APPRPAQLSR	YESQKRRDWN	43	

ФИГ. 10 (1 из 4)

SEQ_ID_NO_210	TFGQYLKNHR	PPLALPRCSG	AHVVEFLHYL	DQFGKTKVHV	PACPFFGHPH	103
SEQ_ID_NO_209	TFMQYLKSON	PPLMMSQFDY	THVLSFLRYL	DQFGKTKVHH	QACVFFGQPD	97
SEQ_ID_NO_230	TFGQYLKNQR	PPVPLSQCNY	NHVLEFLRYL	DQFGKTKVHL	HGCPFFGQPE	111
SEQ_ID_NO_214	TFGQYLKNQR	PPVSLSQSNC	NHVLDLFLRYL	DQFGKTKVHL	HGCVFFGQPN	91
SEQ_ID_NO_239	TFGQYLKNQR	PPVSLAQCSG	NHVLEFLRYL	DQFGKTKVHL	HGCVFFGQPD	89
SEQ_ID_NO_212	TFGQYLKNQR	PPVPLSQCNC	NHVLDLFLRYL	DQFGKTKVHL	QGCIFYGQPE	89
SEQ_ID_NO_226	TFGQYLRNQR	PPVALSQCSS	NHVLEFLRYL	DQFGKTKVHS	QGCLFFGQTE	107
SEQ_ID_NO_220	TFLQYLKNHK	PPLALSRCSG	AHVI EFLKYL	DQFGKTKVHV	ATCPYFGHQQ	92
SEQ_ID_NO_236	TFLQYLRNHK	PPLTLARCSG	AHVI EFLKYL	DQFGKTKVHA	DGCAYFGQPN	131
SEQ_ID_NO_238	TFLQYLRNHK	PPLTLARCSG	AHVI EFLRYL	DQFGKTKVHA	EGCAYFGQPN	145
SEQ_ID_NO_232	TFLQYLRNHR	PPLTLARCSG	AHVI EFLRYL	DQFGKTKVHA	SGCAFYGGPS	91
SEQ_ID_NO_242	TFLQYLRNHR	PPLTLARCSG	AHVI EFLRYL	DQFGKTKVHA	SGCAFYGGPS	91
SEQ_ID_NO_241	TFLQYLRNHR	PPLTLARCSG	AHVI EFLRYL	DQFGKTKVHA	AGCAYYGGPA	94
SEQ_ID_NO_234	TFLQYLRNHR	PPLTLARCSG	AHVI EFLRYL	DQFGKTKVHA	AGCAYYGGPA	93

SEQ_ID_NO_210	PPAPCACPLR	QAWGSLDALI	GRL RAAYEEN	- GGPESNPF	GARAVRLYL R	152
SEQ_ID_NO_209	PPGPCTCPLK	QAWGSLDALI	GRL RAAYEEH	GGSPDTNPF	ANGSI RVHLR	147
SEQ_ID_NO_230	PPGPCTCPLR	QAWGSLDALI	GRL RAAYEEN	- GGLPENNPF	ASGAI RVYLR	160
SEQ_ID_NO_214	PPAPCTCPLR	QAWGSLDALI	GRL RAAYEEH	- GGS AETNPF	GNGAI RVYLR	140
SEQ_ID_NO_239	PPAPCTCPLR	QAWGSLDALI	GRL RAAYEEH	- GGSPETNPF	GNGAI RVYLR	138
SEQ_ID_NO_212	PPAPCTCPLR	QAWGSLDALI	GRL RAAYEEN	- GGSPETNPF	ASGAI RVYLR	138
SEQ_ID_NO_226	PPGPCTCPLR	QAWGSLDALI	GRL RAAYEEN	- GGLPETNPF	ASGTI RVYLR	156
SEQ_ID_NO_220	PPSPCACPLK	QAWGSLDALI	GRL RAAYEEH	- GGRPDSNPF	AARAVRI YLR	141
SEQ_ID_NO_236	PPAPCACPLR	QAWGSLDALI	GRL RAAYEES	- GGRPESNPF	AARAVRI YLR	180
SEQ_ID_NO_238	PPAPCTCPLR	QAWGSLDALI	GRL RAAYEES	- GGRPESNPF	AAKAVRI YLR	194
SEQ_ID_NO_232	PPGPCPCPLR	QAWGSLDALI	GRL RAAYEES	- GGT PESNPF	AARAVRI YLR	140
SEQ_ID_NO_242	PPGPCPCPLR	QAWGSLDALI	GRL RAAYEES	- GGT PESNPF	AARAVRI YLR	140
SEQ_ID_NO_241	PPGPCPCPLR	QAWGSLDALI	GRL RAAYEES	- GGT PESNPF	AARAVRI YLR	143
SEQ_ID_NO_234	PPGPCPCPLR	QAWGSLDALI	GRL RAAYEES	- GGT PESNPF	AARAVRI YLR	142

ФИГ. 10 (2 из 4)

SEQ_ID_NO_210	EVRESQAKAR	GI AYEKRNAS	APQFNSNNSS	SNNMYSSTST	STSTSTS---	199
SEQ_ID_NO_209	EVRESQAKAR	GI PYRKKRR	KTKNVVV--	-----	-----	175
SEQ_ID_NO_230	EVRDFQAKAR	GI CYKKKKKK	RKMNKPT--	-----	-----	188
SEQ_ID_NO_214	EVKECQAKAR	GI PYKKKT	KTQIRARN--	-----	-----	168
SEQ_ID_NO_239	EVRDCQSKAR	GI PYKKKKKK	KNQIKPNN--	-----	-----	166
SEQ_ID_NO_212	EVRECQAKAR	GI PYKKKKKK	PHQKK--	-----	-----	163
SEQ_ID_NO_226	EVRDSQAKAR	GI PYKKKKKK	RNVIR--	-----	-----	181
SEQ_ID_NO_220	EVRESQAKAR	GI PYEKKRK	RAPITVTTA-	--RIDVAPSR	QSEGGGG--	184
SEQ_ID_NO_236	EVREAQAKAR	GI PYEKKRK	RGSTSAPAAP	PPVVTAAEAA	ATSGGGEDDE	230
SEQ_ID_NO_238	DVREAQAKAR	GI PYEKKRK	RGSAAAPPVA	PPPVMTAEAA	GTSSGAGGGD	244
SEQ_ID_NO_232	EVRDSQAKAR	GI PYEKKRK	RSQAAQPAGV	EPGSSSSAAA	AAAGGGD--	187
SEQ_ID_NO_242	EVRDSQAKAR	GI PYEKKRK	RSQAAQPAGV	EPGSSSSAAA	AAAGGGD--	187
SEQ_ID_NO_241	EVRDSQAKAR	GI PYEKKRK	RAQQAAP-	----SSSSSA	APAGGSG--	184
SEQ_ID_NO_234	EVRDSQAKAR	GI PYEKKRK	RAQQQAAAA	ADPASTSSSA	AAAGGSG--	189

SEQ_ID_NO_210	-----	-----	-----	----TSSSS	TSTSSKLRLT	215
SEQ_ID_NO_209	-----	-----	-----	-----	-----VKKDV	180
SEQ_ID_NO_230	-----	-----	-----	-----	-----SSNAHE	194
SEQ_ID_NO_214	-----	-----	-----	-----	-----	168
SEQ_ID_NO_239	-----	-----	-----	-----	-----	166
SEQ_ID_NO_212	-----	-----	-----	-----	-----G	164
SEQ_ID_NO_226	-----	-----	-----	-----	-----PN	183
SEQ_ID_NO_220	-----	-----	-----	-----	----CNDSDPS	191
SEQ_ID_NO_236	DE-----P	SQSAEQQHTT	PASPPTRISS	AGATSTTAAA	AATTTTTTTR	273
SEQ_ID_NO_238	DEDDDEPPPS	ADEPRRQQT	APTPPPASM	ISSTSSSSSS	VAPVTTTTTT	294
SEQ_ID_NO_232	-----	-----	-----	----AGSGGG	AAATTTAQPG	203
SEQ_ID_NO_242	-----	-----	-----	----TSGSGG	AAATTTAQPG	203
SEQ_ID_NO_241	-----	-----	-----	----SISRA	AATAAHQAG	199
SEQ_ID_NO_234	-----	-----	-----	----TSGRAAA	AAAASAAQAG	206

ФИГ. 10 (3 из 4)

SEQ_ID_NO_210	LCP	T	L	P	L	Q	M	S	L	226
SEQ_ID_NO_209	A	N	S	S	T	P	N	Q	S	191
SEQ_ID_NO_230	P	T	T	T	T	F	Q	F	Q	205
SEQ_ID_NO_214	-	-	E	A	K	P	P	M	Q	177
SEQ_ID_NO_239	-	-	E	A	K	S	S	K	Q	175
SEQ_ID_NO_212	G	D	E	S	S	T	I	H	F	175
SEQ_ID_NO_226	G	D	T	S	S	N	L	P	M	194
SEQ_ID_NO_220	V	A	E	A	M	P	P	-	-	198
SEQ_ID_NO_236	G	K	E	K	E	A	A	E	G	284
SEQ_ID_NO_238	S	K	E	K	E	G	S	A	P	305
SEQ_ID_NO_232	G	S	G	T	A	P	S	A	S	212
SEQ_ID_NO_242	G	S	G	T	A	P	S	A	S	212
SEQ_ID_NO_241	G	S	S	A	A	P	S	N	T	208
SEQ_ID_NO_234	G	S	S	A	A	P	S	T	T	215

ФИГ. 10 (4 из 4)

SEQ_ID_NO_437	- MGVEEGEGSM	ETALLEEGSI	SSVQTGKASE	DEYTG DGSD	FWGRPSVKEN	49
SEQ_ID_NO_440	-----M	GSVEDDSSRL	EEALI QDEES	KL YTG DGSD	FKGRPVL KKN	41
SEQ_ID_NO_438	-----	MGSLEERLLV	EDEPVKNDGI	GL YTG DGSD	LDGKPVLKEN	40
SEQ_ID_NO_439	-----	---MKYLF SK	NGGLLEDENS	GL YTR DGSD	IKGNPVLKSE	37
SEQ_ID_NO_435	-----	-----	---MGEVAA	ELYTQDGTID	IKGNPALKKD	26
SEQ_ID_NO_436	-----	-----	-MNAVGGAPD	SEY PQDGTVD	LRGNPVLKAN	29
SEQ_ID_NO_442	-----MKMRA	LVVEGGEMSA	PTPTPKQDKC	CEYTL DGSD	IKGRPAVKGK	45
SEQ_ID_NO_447	MSANEGDLKM	RVIAMGG EAA	SERRAAEEKL	CEYTL DGSD	IKGRPAVKGK	50
SEQ_ID_NO_444	-----	-MEPGI CNSV	NSVNVEATEK	KVTGGNGSYT	EQSDTSYKKI	39
SEQ_ID_NO_431	-----	-----M	APTAVDSKRI	SDIT EDGSD	RRGNPAVKAK	31
SEQ_ID_NO_446	-----	-----M	DATTVD SKW	SPAT EDGSD	RRGNPAVKTT	31
SEQ_ID_NO_433	-----	-----M	DPTTVESKWW	SPTT EDGSD	RRGNPAKAA	31
SEQ_ID_NO_426	-----	-----	--MDQKVRF	EVCTQ DGSD	RHGNPAI RAN	28
SEQ_ID_NO_448	-----	-----	--MDQKVRHT	EVCTQ DGSD	RHGNPAI RAK	28
SEQ_ID_NO_430	-----	-----	---MVN KEL	EVCTLDGSI D	SHGHPAVRER	26
SEQ_ID_NO_428	-----MA	CLNICKEEGS	VQGKLVKKEQ	EVCTR DGSD	RHGDPAI RGR	42
SEQ_ID_NO_429	-----	-----	---MYLNPL	SECTFDGSD	RHGHPAVRAR	26

SEQ_ID_NO_437	TGNWRACPF I	LGNECCERLA	YYGI NTNLVT	- YLTKRLHQQ	NAAAAKSVTT	98
SEQ_ID_NO_440	TGNWKACPF I	LGNECCERLA	YYGI ATNLVK	PI L LAKLHEG	NVSAARNVTT	91
SEQ_ID_NO_438	TGKWRACPF I	LGTECCERLA	YYGI ATNLVT	- YLT SKLHEG	NASAARNVTT	89
SEQ_ID_NO_439	TGNWRACPF I	LGNECCERLA	YYGI AANLVT	- YLT KKLHEG	NVSAARNVTT	86
SEQ_ID_NO_435	TGNWRACPYI	LANECCERLA	YYGMSTNLVN	- FMKDRMGMA	NAAAANNVTN	75
SEQ_ID_NO_436	TGGWKACPYI	LGNECCERLA	YYGI ATNLVT	- YL SHE LHQN	PSTAANNVTN	78
SEQ_ID_NO_442	SGGWL AGALI	LVNQGLATLA	FFGVNVNLVL	- FLTRVLGQS	NEDAANNVSK	94
SEQ_ID_NO_447	SGGWL AGGLI	LVNQGLATMA	FFGVNVNLVL	- FLTRLVQQS	NGDAANNVSK	99
SEQ_ID_NO_444	KGGWKTAI I L	LANQA L ATLA	FFGVGVNLVL	- FLTRVL RQD	SAEAANNVSK	88
SEQ_ID_NO_431	TGNWRS S I LL	LVNYGLVTCA	FFGVGVNLVV	- FLRRVLHQD	NAEAANSI SK	80
SEQ_ID_NO_446	TGRWRS A I LL	LANYGLATCA	FFGVGVNLVV	- FLRRVLHQQ	NAEAANSI SK	80
SEQ_ID_NO_433	SGRWRS A I LL	LANYGLVTCA	FFGVGVNLVV	- FLRRVLHQD	NAEAANNI SK	80
SEQ_ID_NO_426	TGKWL TAI LI	LVNQGLATLA	FFGVGVNLVL	- FLTRVMGQD	NAEAANNVSK	77
SEQ_ID_NO_448	TGKWL TAI LI	LVNQGLATLA	FFGVGVNLVL	- FLTRVMGQD	NAEAANNVSK	77
SEQ_ID_NO_430	TGTWFAGI LI	LVNQGLATLA	FFGVGVNLVL	- FLTRVMGQD	NADAANNVSK	75
SEQ_ID_NO_428	TGTWFAGI LI	LVNQGLATLA	FFGVGVNLVL	- FLTRVLGQD	NAEAANNVSK	91
SEQ_ID_NO_429	TGNW TAI LI	LVNQGLATLA	FFGVGVNLVL	- FLTRVLGQD	NATAANNVSK	75

ФИГ. 11 (1 из 7)

SEQ_ID_NO_437	WAGTCYLTP	F	GAVLADAYW	GRYWTIAAFS	TIYFIGMATL	TLASVSSLK	148
SEQ_ID_NO_440	WQGTCYLAPL	I	GAVLADSYW	GRYWTIAIFS	MIYFIGMGTL	TLASIPALK	141
SEQ_ID_NO_438	WSGTCYLAPL	I	GAVLADSYW	GRYWTIAVFS	IIYFIGMGTL	TLASVAAFK	139
SEQ_ID_NO_439	WQGTCTYI TPL	I	GAVLADAYW	GRYWTIAIFS	TIYFIGMTL	TLASVPAFK	136
SEQ_ID_NO_435	WQGTCTYI TPL	I	GAFLADAYL	GRFWTIA SFM	IIYI FGLGLL	TMATSVHGLV	125
SEQ_ID_NO_436	WSGTCYI TTL	I	GAFLADAYL	GRFWTI VVFS	IIYFLGMVLL	TLSAALPSLK	128
SEQ_ID_NO_442	WTGTVYMFSL	I	GAFLSDSYW	GRYKTCAI FQ	AI FVLGLALL	SVSSHLYLIR	144
SEQ_ID_NO_447	WTGTVYMFSL	I	GAFLSDSYW	GRYKTCAI FQ	AI FVLGLGLL	SLSSRLYLIR	149
SEQ_ID_NO_444	WTGTVYI FSL	I	GAFLSDSYW	GRYLTCAI FQ	FI FVVGLGML	SLSSWRFLIK	138
SEQ_ID_NO_431	WTGTVYI FSL	I	GAFMSDSYW	GRYITCAI FQ	MI YVTGLVIL	SLASWFLLVK	130
SEQ_ID_NO_446	WTGTVYI FSL	I	GAFLSDSYW	GRYVTCAVFQ	IIYVMGLVVL	SLASWLLLVK	130
SEQ_ID_NO_433	WTGTVYI FSL	I	GAFLSDSYW	GRYVTCAI FQ	IIYVTGLVIL	SLASWFLLVK	130
SEQ_ID_NO_426	WTGTVYI FSL	L	GAFLSDSYW	GRYKTCAI FQ	ASFVAGLMML	SLSTGALLLE	127
SEQ_ID_NO_448	WTGTVYI FSL	L	GAFLSDSYW	GRYKTCAI FQ	ASFVAGLVML	SLSTGALLLE	127
SEQ_ID_NO_430	WTGTVYI FSL	L	GAFLSDSYW	GRYITCAI FQ	VI FVVGLVAL	SLTSYIFLLK	125
SEQ_ID_NO_428	WTGTVYI FSL	L	GAFLSDSYW	GRYKTCAI FQ	AI FVTGLVLL	SLSSYFLLK	141
SEQ_ID_NO_429	WTGTVYI FSL	L	GAFLSDSYW	GRYKTCAI FQ	VI FVI GLVLL	SLSSYIFLLK	125

SEQ_ID_NO_437	PPSC- - I GSD	CPTANLAQYG	VFFLGLYLI A	LGTGGI KPCV	SSFGADQFDD	196
SEQ_ID_NO_440	PAEC- - LGAV	CPPATPAQYA	VFFI GLYLI A	LGTGGI KPCV	SSFGADQFDD	189
SEQ_ID_NO_438	PSPC- - VGSV	CPAATPAQYA	VFFGLYLI A	LGTGGI KPCV	SSFGADQFDD	187
SEQ_ID_NO_439	PPQC- - VGSV	CPSASPAQYA	IFFFGLYLI A	LGTGGI KPCV	SSFGADQFDD	184
SEQ_ID_NO_435	PALCASKGV-	CDPTPGQSA	AVFI ALYLI A	LGTGGI KPCV	SSFGADQFDE	172
SEQ_ID_NO_436	PPSGEGVVAL	- - - - SSTQLA	VFYLA LYLI A	LGTGGI KPCV	SSFGADQFDE	173
SEQ_ID_NO_442	PDGCGMEHAP	CGPHSGKELG	IFYI ALYMI A	FGNGGYQPNI	ATFGADQFDE	194
SEQ_ID_NO_447	PVGCGTEHTP	CASHSGTEMG	IFYI ALYMI A	FGNGGYQPNI	ATFGADQFDE	199
SEQ_ID_NO_444	PVGCGNEETT	CLEPSSVGVG	IFYLSI YLVA	FGYGGHQPTL	ATFGADQFDE	188
SEQ_ID_NO_431	PTGCGAAGEH	CDAPSSAGVA	LFYLS TYMI A	FGNGGYQPSI	ATFGSDQFDE	180
SEQ_ID_NO_446	PSGCGGVKAH	CDGPSAPGVA	LFYLS TYMI A	FGNGGYQPSI	ATFGSDQFDE	180
SEQ_ID_NO_433	PSGCGGV DAR	CDEPSAPGVA	LFYLS TYMI A	FGNGGYQPSI	ATFGSDQFDE	180
SEQ_ID_NO_426	PSGCGVEDSP	CKPHSTFKTV	LFYLSVYLI A	LG YGGYQPNI	ATFGADQFDA	177
SEQ_ID_NO_448	PSGCGVEESP	CKPHSTVKTV	IFYLSVYLI A	LG YGGYQPNI	ATFGSDQFDA	177
SEQ_ID_NO_430	PNGCGSKEIP	CGTHSSYETT	LFYVSI YLVA	LGNGGYQPTI	ATFGADQFDE	175
SEQ_ID_NO_428	PRGCGDEHSP	CGSHSTYQNV	FFYFSI YLVA	LGNGGYQPNI	ATFGADQFDE	191
SEQ_ID_NO_429	PNGCGDKFEP	CGSHSTFEIS	FFYLSI YLI A	LGNGGYQPNI	ATFGADQFDE	175

SEQ_ID_NO_437	TDPKEKKKKG	SFFNWFYFSI	NVGALVSSSV	L VWQDNVGV	GWGFGI PTLF	246
SEQ_ID_NO_440	TDSRERVKKG	SFFNWFYFSI	NI GALI SSSF	I VWI QENAGW	GL GFGI PALF	239
SEQ_ID_NO_438	TDPVERVQKG	SFFNWFYFSI	NI GALI SSSF	L VWI QDNAGW	GL GFGI PTLF	237
SEQ_ID_NO_439	TDPKERVKKG	SFFNWFYFSI	NI GALI SSSL	I VWI QENAGW	GL GFGI PAVF	234
SEQ_ID_NO_435	HDDVERKSKS	SFFNWFYFSI	NI GALVASSV	L VYVQTHVGV	SWGFGI PAVV	222
SEQ_ID_NO_436	NDVKEKKRKS	SFFNWFYFTI	NI GALI ASSA	L VYI QENVGV	GWGFGI PAVA	223
SEQ_ID_NO_442	DDPAEAHSKV	SFFSYFYLLAL	NLGSLFSNTF	L SYLEDKGSW	AL GF WASTAA	244
SEQ_ID_NO_447	EDPAEAHSKV	SFFSYFYMAL	NLGSLFSNTF	L SYLQDHGKW	VL GF WASTAA	249
SEQ_ID_NO_444	KNEKQKNARE	AFFSYFYFAL	NVGSLFSNTI	L VYYEDSGMW	TMGF LVSLAS	238
SEQ_ID_NO_431	TDPREARSKV	AFFSYFYLLAL	NVGSLFSNTV	L VYYEDEGRW	VMGF WWSAAA	230
SEQ_ID_NO_446	TDPEEARSKV	AFFSYFYLLAL	NVGSLFSNTV	L VYYEDSGRW	VMGF WWSAAA	230
SEQ_ID_NO_433	TDPKEARSKV	AFFSYFYLLAL	NVGSLFSNTV	L VYYEDSGRW	VMGF WWSAAA	230
SEQ_ID_NO_426	EDSVEGHSKI	AFFSYFYLLAL	NLGSLFSNTV	L GYFEDQGEW	PL GF WASAGS	227
SEQ_ID_NO_448	DDSVEGHSKI	AFFSYFYLLAL	NLGSLFSNTV	L GYFEDQGAW	PL GF WASAGS	227
SEQ_ID_NO_430	SDPSEQHSKI	AFFSYFYLLAL	NI GSLFSNTI	L DYFEDDGLW	TL GF CVSAGS	225
SEQ_ID_NO_428	EDPKEGHSKI	AFFSYFYLLAL	NLGSLFSNTI	L GYFEDRGMW	AL GF WASAGS	241
SEQ_ID_NO_429	DNPKESSHVKV	AFFSYFYLLAL	NLGSLFSNTI	L GYFEDGGRW	VL GF WASSAS	225
SEQ_ID_NO_437	MGLAI GSFFS	GTPLYRLQKP	GGSPVTRMCQ	VVASLRKLR	KT VPLDHS	295
SEQ_ID_NO_440	MGLAI GSFFL	GTPLYRFQKP	GGSPLTRMCQ	VVAASFRKR	LT VPEDSSL	288
SEQ_ID_NO_438	MGVAI I SFFS	GTPLYRFQKP	GGSPLTRMCQ	VFVASFRKWN	LD VPQDSSL	286
SEQ_ID_NO_439	MGI AI ASFFF	GTPLYRFQKP	GGSPLTRMCQ	VLVAVFHKWN	LS VPDDSTL	283
SEQ_ID_NO_435	MAI AVGSFFV	GTPLYRHQRP	GGSPLTRI AQ	VLVAATRKL	- - VPVDGSA	269
SEQ_ID_NO_436	MGI AI VSFLI	GSPLYRHQKP	GGSPI TRI AQ	VLVAATRKL	MK VPNGKH	272
SEQ_ID_NO_442	AATALLLFLS	GT LRYRYFQP	GGNPI GRI CQ	VAI AASRKWK	AGASTTGVVS	294
SEQ_ID_NO_447	AATALLLFLS	GTPQYRHAQP	CGNPMASI CQ	VASAAQRNWK	SGGVSPDVEI	299
SEQ_ID_NO_444	AVI ALVSYLA	G YRKYRYVK	YGNPVI RVVQ	VFVA TVRKWK	VG - PAKEHQ	286
SEQ_ID_NO_431	AAMALVLFLL	GTPNYRHF KP	TGNPLTRI AQ	VFVA AFRKWR	AE VPRSEL	278
SEQ_ID_NO_446	AALALVLFLL	GTPGYRHF KP	SGNPLTRVAQ	VFVAALRKWR	AE VPRGEL	278
SEQ_ID_NO_433	AALALVLFLL	GTPNYRHF KP	SGNPLTRVAQ	VFVAALRKWH	AE VPREEF	278
SEQ_ID_NO_426	AFAGLVLFLL	GTPKYRHFTP	RESPWSRFCQ	VLVAATRKA	I D VHEELN	276
SEQ_ID_NO_448	AFAGLVLFLL	GTPKYRHFTP	RESPWSRFCQ	VLVAATRKA	I D VNYDDMN	276
SEQ_ID_NO_430	AALALVLFLL	GT SKYRYFKP	VGNPLPRFCQ	VFVAALRKWK	VQ MF DGEDK	274
SEQ_ID_NO_428	ALLALVLFLL	GTPRYRHFTP	KGNPLSRCCQ	VMAATRCKWK	VQRMPNQGDD	291
SEQ_ID_NO_429	ALLALVLFLL	GTPRYRHFTP	SGNPLSRFCR	VVVAATRCKWK	VE MPPEGED	274

ФИГ. 11 (3 из 7)

SEQ_ID_NO_437	LYEV	- Q-	DR	NSAI	QGS	RKL	EHT	DEFK	CLD	KSAI	I TDD	DDV	- - - -	RKGG	F S	338
SEQ_ID_NO_440	LYET	- P-	DK	SSAI	EGS	RKL	QHS	DEL	RCLD	RAAVI	SDDER	- - - -	KRGD	YS	331	
SEQ_ID_NO_438	LFEL	- P-	DK	TSAI	EGS	RKL	EHS	DEL	RCLD	KAAV	VSD	DV	- - - -	QQGD	L T	329
SEQ_ID_NO_439	LYET	- P-	DK	SSAI	EGS	RKL	LHT	DEL	RCLD	KAAV	VSDNEL	- - - -	TTGD	YS	326	
SEQ_ID_NO_435	LYET	- A-	DR	ESGI	EGS	RKL	EHT	EQFR	FLD	KAAV	ETQAD	- - - -	KTAT	GP	311	
SEQ_ID_NO_436	LYEA	- D-	DK	ESGI	EGS	RKL	EHT	EEFR	FLD	KAAI	PRGDEE	- - - -	LQGT	RP	315	
SEQ_ID_NO_442	LYEG	- - -	DE	KADA	AAG	RKL	LHT	QGF	SFLD	RAA	HADTD	- - - -	SKL	GAR	334	
SEQ_ID_NO_447	LYE	- - -	GD	DKTD	SGS	RKL	LHT	KGFR	FLD	RAAL	T TEDTN	- - - -	SKLAT	CSKTR	344	
SEQ_ID_NO_444	LYEV	- D-	GP	ESAI	KGS	RKI	HHS	NDFR	FMD	KAAT	I TEKDA	- - - -	VNL	K	327	
SEQ_ID_NO_431	LHEV	- D-	GD	ESQI	AGI	RKI	LHS	DQ	RFLD	KAAT	VTEEDY	C- - -	TPEN	MQ	322	
SEQ_ID_NO_446	LHEV	VEA-	GE	DPKV	SGI	RKI	LHS	DGL	RFLD	KAAT	I TEEEE	- - - -	EEG	- -	320	
SEQ_ID_NO_433	LHEV	- E-	GE	DPKV	SGI	RKI	LHS	DEL	RFLD	RAAT	VTEEEY	G- - -	APEK	MK	322	
SEQ_ID_NO_426	LYDS	- - - -	-	ETQY	TGDK	KI	LHT	KGFR	FLD	RAAI	VTPDDE	- - - -	AEKV	ESGSKY	320	
SEQ_ID_NO_448	LYDS	- - - -	-	ETQRT	TGDK	KI	LHT	KGFR	FLD	RAAI	VTPDDE	- - - -	AEKV	ESGSAY	320	
SEQ_ID_NO_430	LHEV	- - -	E	DCL	SNG	RKM	YHT	QGF	RFLD	KAA	F I TPKDL	KQ- -	MEEN	KC	317	
SEQ_ID_NO_428	QFES	- DVP	PK	DGSK	NGDR	KI	LHT	QGF	RFLD	RAAI	I TSKDY	- - - -	TENR	LH	335	
SEQ_ID_NO_429	LYEV	- D-	GD	DCSG	NCR	RKM	LHT	QGF	KFLD	KAAV	LTA	KER	EQ- -	EDKE	SR	319

SEQ_ID_NO_437	NPWRL	CTVTQ	VEEM	KI L	LRM	FPI	WATGI	VF	AAVYS	QI STM	FVEQ	GMTLD	- - - -	- - - -	387
SEQ_ID_NO_440	NLWRL	CTVTQ	VEEL	KI L	IRM	FPV	WATGI	VF	SAVYA	QIMSTM	FVEQ	GTMMD	- - - -	- - - -	380
SEQ_ID_NO_438	NPWRL	CTVTQ	VEEL	KI L	IRM	FPI	WATGI	VF	SAVYA	QIMSTM	FVEQ	GTMVLD	- - - -	- - - -	378
SEQ_ID_NO_439	NAWRL	CTVTQ	VEEL	KI L	IRM	FPI	WATGI	VF	SAVYA	QIMSTM	FVEQ	GTMVMD	- - - -	- - - -	375
SEQ_ID_NO_435	SPWRL	CTVTQ	VEEL	KSV	VRL	LPI	WASGI	VF	ATVY	QIMSTM	FV	QGN	TLD	- - - -	360
SEQ_ID_NO_436	SGWRL	TSVTQ	VEEV	KI V	MRL	LPI	WASGI	VF	ATVYS	QIMSTM	FV	QGAL	LMN	- - - -	364
SEQ_ID_NO_442	DPWKL	CTVTQ	VEEV	KSI L	LRL	LPI	WLCTI	LY	SVVFT	QMASL	FVVQ	GAMRR	- - - -	- - - -	384
SEQ_ID_NO_447	DQWRL	CTVTQ	VEQV	KSI L	RI	LPI	WWCTI	LY	SVVFT	QMASL	FVVQ	GAMRR	- - - -	- - - -	394
SEQ_ID_NO_444	NHWRL	CTVTQ	VEE	A KCV	LRM	LPV	WLCTI	I Y	SVVFT	QMASL	FVEQ	GDMN	- - - -	- - - -	376
SEQ_ID_NO_431	DPWRL	CTVTQ	VEEV	KCI L	KM	LPI	WLCTI	VY	SVVFT	QMASL	FVEQ	GTTMN	- - - -	- - - -	371
SEQ_ID_NO_446	- PWRL	CTVTQ	VEEV	KCI L	RM	LPI	WLCTI	VY	SVVFT	QMASL	FVEQ	GATMD	- - - -	- - - -	368
SEQ_ID_NO_433	DPWRL	CTVTQ	VEEV	KCI L	KM	LPI	WLCTI	VY	SVVFT	QMASL	FVEQ	GATMN	- - - -	- - - -	371
SEQ_ID_NO_426	DPWRL	CSVTQ	VEEV	KCVL	LRL	LPI	WLCTI	LY	SVVFT	QMASL	FVVQ	GAMK	- - - -	- - - -	369
SEQ_ID_NO_448	DPWRL	CSVTQ	VEEV	KCVL	LRL	LPI	WLCTI	LY	SVVFT	QMASL	FVVQ	GAMK	- - - -	- - - -	369
SEQ_ID_NO_430	SPWF	LSVTQ	VEEV	KCI L	LRL	LPI	WLCTI	LF	SVVFS	QMASL	FVEQ	GAMK	- - - -	- - - -	366
SEQ_ID_NO_428	DPWRV	CTVNQ	VEE	F KCI	LRL	LPI	WLCTI	LY	SVVFT	QMASL	FVEQ	GADMK	- - - -	- - - -	384
SEQ_ID_NO_429	SPWRI	CTVTQ	VEEV	KCI L	LRL	LPI	WLCTI	LY	SVVFT	QMASL	FVEQ	GAMK	- - - -	- - - -	368

ФИГ. 11 (4 из 7)

SEQ_ID_NO_437	- TSI GSRFHI	PPASLSVFDV	VSVM L WWPVY	DRVI VPI VSK	YTK RE RGF	434
SEQ_ID_NO_440	- TSVGS FKI	PAASLSTFDV	I SVI FWWPVY	DRFI VPI ARK	FTG KE RGF	426
SEQ_ID_NO_438	- TTI GS F T	PPASLSTFDV	LSVI VWWPMY	DRL LVPLARK	FTG KE RGF	424
SEQ_ID_NO_439	- TAVGS FKI	PAASLSTFDT	I SVI VWWPVY	DKI LVPI ARR	FTG IE RGF	421
SEQ_ID_NO_435	- ASMGP KFKI	PSASLSI FDT	LSVI AWWPVY	DRI LVP AVRS	VTG RP RGF	407
SEQ_ID_NO_436	- VSMGK ANI	PSASLSI FDT	I SVI VCV M Y	DRFLVPVVRK	RTG HV RGF	410
SEQ_ID_NO_442	TTPE SG FSV	PPSSMSAFDI	LAVATTI FLY	RRAI CP ELAR	LTG RP AGP	431
SEQ_ID_NO_447	TTPLGS FSI	PASSMSAFDI	LTVT TTI FLY	RRAI CPLLAR	LTG RP TGP	441
SEQ_ID_NO_444	- NKI GN FHL	PAASMSVFDI	CSVL L CTG Y	RQI LVPLAGR	FSG NP RGL	422
SEQ_ID_NO_431	- TNI GS FHV	PAASMSVFDI	LSVLAFI AI Y	RRVLVPVMSR	LSG NP QGL	417
SEQ_ID_NO_446	- TNI GS FHF	PAASMSLFDV	LSVLAFI AI Y	RRVLVPVMAR	LSG NP QGL	414
SEQ_ID_NO_433	- TNI GS FHF	PAASMSLFDI	LSVLAFI AI Y	RRVLVPVMAR	LSG NP QGL	417
SEQ_ID_NO_426	- TNI KN FRI	PASSMSSFDI	LSVAFFI FAY	RRFLDPLFAR	LNKTER NKGL	417
SEQ_ID_NO_448	- TNI KD FRI	PASSMSTFDI	LSVAFFI FAY	RRFLDPLFAR	LNKTEPNKGL	417
SEQ_ID_NO_430	- TKI ST FHI	PPASMSSFDI	LSVVSFI FI Y	RRI LDPLVAR	FTK KS KGL	412
SEQ_ID_NO_428	- TTI SK FHI	PPASMSSFDI	LSVAAFI FI Y	RRVL DPLVAR	IRK DP KGL	430
SEQ_ID_NO_429	- TTL LG FHI	PPASMSSFDI	VSVAAFI FI Y	RRVL DPLVAR	LKGTKA RGL	415

SEQ_ID_NO_437	TELQRMGI GL	F I S I LAMVAA	ALVEI RRL N	V KTYDLVYD	KGTPV - - - P	479
SEQ_ID_NO_440	SELQRMGI GL	F I SVLCMSAA	AI VEI KRLQL	A - KELDLVD	KAVPV - - - P	470
SEQ_ID_NO_438	SELQRMGI GL	F I S I L S M TAA	AI VEI RRLQI	A - K S L GLVD	QNVAV - - - P	468
SEQ_ID_NO_439	SELQRMGI GL	FLSMLCMSAA	AI VEI RRLQL	A - RDLGLVD	EAVSV - - - P	465
SEQ_ID_NO_435	TQLQRMGI GL	VVSM E AMLAA	GVLEL VRLRT	I - AQRGLYG	EHDV - - - P	451
SEQ_ID_NO_436	TQLQRMGI GL	F I SVLAMVVA	AI VEI ERLKL	A - RRDGVAGN	PQDEALPVES	459
SEQ_ID_NO_442	TELQRMGL GL	VVGALAMATA	GTVEHFRKAR	A - - - - - T	AAMSS - - - D	469
SEQ_ID_NO_447	TELQRMGL GL	VLGAMAMANA	GTVEHFRKAS	A - - - - - T	TANGS - - - E	479
SEQ_ID_NO_444	TELQRMGVGL	I I GMLAI LAA	GAT EFERLKH	I - - - - - TP	GEKAS - - - S	461
SEQ_ID_NO_431	TELQRMGVGL	VVGMAAMVVA	GVVEVERLKR	V - - - - - GA	PDQPS - - - S	456
SEQ_ID_NO_446	TELQRMGVGL	VI GMAAMVVA	GVVEVERLKR	V - - - - - AA	PDQPS - - - S	453
SEQ_ID_NO_433	TELQRMGVGL	VI GMAAMVVA	GAVEVERLRR	V - - - - - AA	PDQPS - - - S	456
SEQ_ID_NO_426	TELQRMGI GL	VI AI MAMI SA	GI VEI HRLKN	K - EPE SATS	ISSSS - - - T	461
SEQ_ID_NO_448	TELQRMGI GL	VI AI TAMI SA	GI VEI YRLKH	K - - - - - ET	ASNSS - - - S	456
SEQ_ID_NO_430	TELQRMGI GL	VLAI I AMVSA	GLVEI FRLKY	AI KEEKNC SH	CEGTS - - - S	458
SEQ_ID_NO_428	TELQRMGI GL	VI AI I AMVSA	GI VELFRLKY	A - - RKDCPR	CESAS - - - S	473
SEQ_ID_NO_429	TELQRMGI GL	I I AI I AMVAA	GI VEM FRLKY	A - - KKDCRQ	CESES - - - S	458

ФИГ. 11 (5 из 7)

SEQ_ID_NO_437	MSI FWQI PQY	FLVGASEVFT	FVGQLEFFYD	QSPDAMRSLC	SALSLLTTAL	529
SEQ_ID_NO_440	LTIFQI PQY	FLLGAAEVFT	FVGQLEFFYD	QSPDAMRSLC	SALSLLTTSL	520
SEQ_ID_NO_438	MSI LWQI PQY	FLVGASEIFT	FIGQLEFFYD	QSPDAMRSLC	SALSLLTTAL	518
SEQ_ID_NO_439	LSI FWQI PQY	FILGAAEIFT	FIGQLEFFYD	QSPDAMRSLC	SALSLLTTAL	515
SEQ_ID_NO_435	ISI FWQVPQY	FII GAAEVFT	FVGQLEFFYD	QAPDAMRSLC	SALSLLTTVAL	501
SEQ_ID_NO_436	LTIFVQI PQY	FLI GAAEVFT	FVGQLEFFYD	QAPDAMRSLM	SALSLLTTVAL	509
SEQ_ID_NO_442	LHI MWQVPQY	ALI GVSEVMM	YVGQLEFFNG	QMPDGLKSFG	SALCMMSMSL	519
SEQ_ID_NO_447	LHI LWQVPQY	ALI GVSEVMM	YVGQLEFFNG	EMPDGEKSFG	SALCMMSMSL	529
SEQ_ID_NO_444	LSI FWQI PQY	VLVGASEVFM	YVGQLEFFNG	QAPDGIKSFG	SSLCMASISL	511
SEQ_ID_NO_431	LSVLWQVPQY	ALI GASEVFM	YVGQLEFFNG	QAPDGVKSFG	SSLCMASISL	506
SEQ_ID_NO_446	LSVLWQVPQY	ALI GASEVFM	YVGQLEFFNG	QAPDGVKSFG	SALCMASISL	503
SEQ_ID_NO_433	LSVLWQVPQY	ALI GASEVFM	YVGQLEFFNG	QAPDGVKSFG	SALCMASISL	506
SEQ_ID_NO_426	LSI FWQVPQY	MLI GASEVFM	YVGQLEFFNS	QAPTGLKSFA	SALCMASISL	511
SEQ_ID_NO_448	LSI FWQVPQY	MMI GASEVFM	YVGQLEFFNS	QAPTGLKSFA	SALCMASISL	506
SEQ_ID_NO_430	LSI FWQVPQY	VLI GASEVFM	YVGQLEFFNS	QAPDGLKSFG	SALCMTSISL	508
SEQ_ID_NO_428	LSI LWQI PQY	VLI GASEVFM	YVGQLEFFNG	QAPDGLKSFG	SALCMTSISL	523
SEQ_ID_NO_429	LSI FWQI PQY	VLI GASEVFM	YVGQLEFFND	QAPDGLKSFG	SALCMTSISL	508

SEQ_ID_NO_437	GNYLSSLIVT	I VMFATTK	G	GNL GW PDNL	NEGHLDYFFW	VLASLSVLNL	578
SEQ_ID_NO_440	GNYLSSFILT	VVLYFTTR	G	GNPGW PDNL	NKGHLDYFSG	LAGLSFLNM	568
SEQ_ID_NO_438	GNYLSSFILT	MVTITTR	G	GKPGW PDNL	NEGHLDYFFW	LAGLSFLNL	567
SEQ_ID_NO_439	GNYLSSFILT	VVTSITTR	G	GKPGW PNNL	NKGHLDYFFW	LLAALSFFNL	564
SEQ_ID_NO_435	GNYLSTLLVT	VVAKLTTR	G	GKQGW PDNL	NVGHLDYFFW	LLAALSLLNF	550
SEQ_ID_NO_436	GNYLSSVLVT	I VTEVTTK	G	GKPGW PNNL	NRGHLDYFFW	MLAILSLNL	558
SEQ_ID_NO_442	GNYFSDVI VS	AVTRLTTT	R	GRSGW PADL	NEGHLDKFYF	LLAVLAVADF	568
SEQ_ID_NO_447	GNYFSDI I VS	AVTKATAV	D	GRPGW PADL	NEGHLNKFYF	LLAILSVADF	578
SEQ_ID_NO_444	GNYVSSLLVY	MVMGITAR	G	ENPGW PNNL	NVGHMDRFFF	LVAVLTALDF	560
SEQ_ID_NO_431	GNYVSI MLVS	VVTSLTAG	D	RRPGW PGNL	NSGHLDRFYF	LLAALS LVDL	555
SEQ_ID_NO_446	GNYVSI MLVS	VVTSLTAG	E	KRPGW PGNL	NSGHLDRFYF	LLAALS LVDL	552
SEQ_ID_NO_433	GNYVSI MLVS	VVTSLTAG	E	RRPGW PGNL	NSGHLDRFYF	LLAALS LVDL	555
SEQ_ID_NO_426	GNYVSSLLVS	I VMKI STT	D	DVHGW PENL	NKGHLERFYF	LLAGLTAADF	560
SEQ_ID_NO_448	GNYVSSLLVS	I VMKI STR	D	DLPGW PGNL	NKGHLDRFYF	LLAALTAADF	555
SEQ_ID_NO_430	GNYVSSLLVA	I VMKI STRNE	G	MLGW PGNL	NMGHLDRFYF	LLAALTAADL	558
SEQ_ID_NO_428	GNYVSSLLVT	VVMKI STR	D	EMPGW PGNL	NKGHLDRFFF	LLAVLTADL	572
SEQ_ID_NO_429	GNYVSSLLVT	I VMKFSTR	D	QMPGW PSNL	NKGHLDRFYF	LLAALTMADF	557

ФИГ. 11 (6 из 7)

SEQ_ID_NO_437	L V Y V S C A H R Y	- K Y K K A S	- - - - -	- - - - -	594
SEQ_ID_NO_440	F L Y I V A A K R Y	- K S K K A S	- - - - -	- - - - -	584
SEQ_ID_NO_438	V I Y V F C A A R Y	- K C K K A S	- - - - -	- - - - -	583
SEQ_ID_NO_439	V I Y V F L C Q M Y	- K S K K A S	- - - - -	- - - - -	580
SEQ_ID_NO_435	A V Y L L I A S W Y	- T Y K K T A G D D	H P D A I A K G G A	H D - - - - -	581
SEQ_ID_NO_436	I F Y L V V A K F Y	- T Y K R V H N A A	D V G E D G K P S K	S G Y E - - - -	591
SEQ_ID_NO_442	A V Y L V C A S R Y	- G S G K V D G R S	S D D E E E G A A G	Q V T S L P A Y A	606
SEQ_ID_NO_447	A V Y L V F A G R Y	R K S C K V E G R S	D D E E E G S V D D	E E A C H A - - -	614
SEQ_ID_NO_444	V L Y L L C A R W Y	- K S I N L G D G D	M G - - - - S Q E D	E E V I - - - S K V	592
SEQ_ID_NO_431	A V Y V A C A V W Y	- K G I K L D S N E	E K - - - - A N K I	T V H V - - - -	584
SEQ_ID_NO_446	A V Y V A C A M W Y	- K G I K L D G G D	G D N - - - R R K V	S A H V - - - -	582
SEQ_ID_NO_433	A V Y V A C A T W Y	- K G I K L D G G G	E - - - - T R K N	P A H V - - - -	583
SEQ_ID_NO_426	V V Y L I C A K W Y	- K Y I K S E A S F	S E - - - - S V T E	E E E V - - - -	589
SEQ_ID_NO_448	V V Y L V C A K W Y	- K Y I K S E A S F	S E - - - - S I A E	E E - - - - -	582
SEQ_ID_NO_430	L V Y I A M A R W Y	- K Y V K F H G N N	I E Q G - - N N I E	E N V V - - - -	589
SEQ_ID_NO_428	V L Y I I C A R W Y	- K Y I K F E R Q Q	E A D S I L S N E E	A D L R V - - -	606
SEQ_ID_NO_429	G V Y I I C A K W Y	- K S I K L E G K Y	E D S N - - S K G S	D D V M V - - -	589

ФИГ. 11 (7 из 7)

ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ
(статья 15(3) ЕАПК и правило 42 Патентной инструкции к ЕАПК)

Номер евразийской заявки:

202091701

А. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ:

C12N 15/82 (2006.01)

A01H 5/00 (2006.01)

Согласно Международной патентной классификации (МПК)

Б. ОБЛАСТЬ ПОИСКА:

Просмотренная документация (система классификации и индексы МПК)

C12N 15/82, A01H 5/00

Электронная база данных, использовавшаяся при поиске (название базы и, если, возможно, используемые поисковые термины)

В. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ

Категория*	Ссылки на документы с указанием, где это возможно, релевантных частей	Относится к пункту №
Y	WO 2006/005751 A1 (CROPDESIGN N.V. et.al.) 19.01.2006, формула, примеры 1, 2	1-13
Y	ALEXANDROV N. N. et.al. Features of Arabidopsis genes and genome discovered using full-length cDNAs, Plant Molecular Biology, 2006, v.60, pp.69-85, весь документ	1-7
Y	WO 2008/150461 A2 (SAPPHIRE ENERGY et al.) 11.12.2008, формула, параграфы [0003]-[0012]	1-13

☐ последующие документы указаны в продолжении

* Особые категории ссылочных документов:

«А» - документ, определяющий общий уровень техники

«D» - документ, приведенный в евразийской заявке

«E» - более ранний документ, но опубликованный на дату подачи евразийской заявки или после нее

«O» - документ, относящийся к устному раскрытию, экспонированию и т.д.

"P" - документ, опубликованный до даты подачи евразийской заявки, но после даты испрашиваемого приоритета"

«Т» - более поздний документ, опубликованный после даты приоритета и приведенный для понимания изобретения

«X» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий новизну или изобретательский уровень, взятый в отдельности

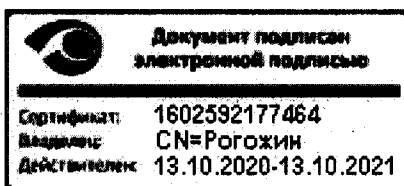
«Y» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий изобретательский уровень в сочетании с другими документами той же категории

«&» - документ, являющийся патентом-аналогом

«L» - документ, приведенный в других целях

Дата проведения патентного поиска: **16/12/2020**

Уполномоченное лицо:
Начальник Управления экспертизы



Д.Ю. Рогожин